

1 Trans-ancestry meta-analyses identify novel rare and common variants associated with blood

2 pressure and hypertension

3 Praveen Surendran ^{1,167}, Fotios Drenos ^{2,3,167}, Robin Young ^{1,167}, Helen Warren ^{4,5,167}, James P Cook ^{6,7,167},
4 Alisa K Manning ^{8,9,10,167}, Niels Grarup ^{11,167}, Xueling Sim ^{12,13,14,167}, Daniel R Barnes ¹, Kate Witkowska ^{4,5},
5 James R Staley ¹, Vinicius Tragante ¹⁵, Taru Tukiainen ^{8,9,16}, Hanieh Yaghootkar ¹⁷, Nicholas Masca ^{18,19},
6 Daniel F Freitag ¹, Teresa Ferreira ²⁰, Olga Giannakopoulou ²¹, Andrew Tinker ^{21,5}, Magdalena Harakalova
7 ¹⁵, Evelin Mihailov ²², Chunyu Liu ²³, Aldi T Kraja ^{24,25}, Sune Fallgaard Nielsen ²⁶, Asif Rasheed ²⁷, Maria
8 Samuel ²⁷, Wei Zhao ²⁸, Lori L Bonnycastle ²⁹, Anne U Jackson ^{13,12}, Narisu Narisu ²⁹, Amy J Swift ²⁹,
9 Lorraine Southam ^{30,20}, Jonathan Marten ³¹, Jeroen R Huyghe ^{13,12}, Alena Stančáková ³², Cristiano Fava ^{33,34},
10 Therese Ohlsson ³³, Angela Matchan ³⁰, Kathleen E Stirrups ^{21,35}, Jette Bork-Jensen ¹¹, Anette P Gjesing ¹¹,
11 Jukka Kontto ³⁶, Markus Perola ^{36,37,22}, Susan Shaw-Hawkins ⁴, Aki S Havulinna ³⁶, He Zhang ³⁸, Louise A
12 Donnelly ³⁹, Christopher J Groves ⁴⁰, N William Rayner ^{40,20,30}, Matt J Neville ^{40,41}, Neil R Robertson ^{20,40},
13 Andrianos M Yiorkas ^{42,43}, Karl-Heinz Herzig ^{44,45}, Eero Kajantie ^{36,46,47}, Weihua Zhang ^{48,49}, Sara M
14 Willems ⁵⁰, Lars Lannfelt ⁵¹, Giovanni Malerba ⁵², Nicole Soranzo ^{53,35,54}, Elisabetta Trabetti ⁵², Niek
15 Verweij ^{55,9,56}, Evangelos Evangelou ^{48,57}, Alireza Moayyeri ^{48,58}, Anne-Claire Vergnaud ⁴⁸, Christopher P
16 Nelson ^{18,19}, Alaitz Poveda ^{59,60}, Tibor V Varga ⁵⁹, Muriel Caslake ⁶¹, Anton JM de Craen ^{62,63}, Stella
17 Trompet ^{62,64}, Jian'an Luan ⁵⁰, Robert A Scott ⁵⁰, Sarah E Harris ^{65,66}, David CM Liewald ^{65,67}, Riccardo
18 Marioni ^{65,66,68}, Cristina Menni ⁶⁹, Aliko-Eleni Farmaki ⁷⁰, Göran Hallmans ⁷¹, Frida Renström ^{59,71}, Jennifer
19 E Huffman ^{31,23}, Maija Hassinen ⁷², Stephen Burgess ¹, Ramachandran S Vasan ^{23,73,74}, Janine F Felix ⁷⁵,
20 CHARGE-Heart Failure Consortium ⁷⁶, Maria Uria-Nickelsen ⁷⁷, Anders Malarstig ⁷⁸, Dermot F Reilly ⁷⁹,
21 Maarten Hoek ⁸⁰, Thomas Vogt ^{80,81}, Honghuang Lin ^{23,82}, Wolfgang Lieb ⁸³, EchoGen Consortium ⁷⁶,
22 Matthew Traylor ⁸⁴, Hugh F Markus ⁸⁴, METASTROKE Consortium ⁷⁶, Heather M Highland ⁸⁵, Anne E
23 Justice ⁸⁵, Eirini Marouli ²¹, GIANT Consortium ⁷⁶, Jaana Lindström ³⁶, Matti Uusitupa ^{86,87}, Pirjo
24 Komulainen ⁷², Timo A Lakka ^{72,88,89}, Rainer Rauramaa ^{72,89}, Ozren Polasek ^{90,91}, Igor Rudan ⁹⁰, Olov
25 Rolandsson ⁹², Paul W Franks ^{59,92,93}, George Dedoussis ⁷⁰, Timothy D Spector ⁶⁹, EPIC-InterAct
26 Consortium ⁷⁶, Pekka Jousilahti ³⁶, Satu Männistö ³⁶, Ian J Deary ^{65,67}, John M Starr ^{65,94}, Claudia
27 Langenberg ⁵⁰, Nick J Wareham ⁵⁰, Morris J Brown ⁴, Anna F Dominiczak ⁹⁵, John M Connell ³⁹, J Wouter
28 Jukema ^{64,96}, Naveed Sattar ⁹⁵, Ian Ford ⁶¹, Chris J Packard ⁶¹, Tõnu Esko ^{22,97,8,9}, Reedik Mägi ²², Andres
29 Metspalu ^{22,98}, Rudolf A de Boer ⁹⁹, Peter van der Meer ⁹⁹, Pim van der Harst ^{99,100,101}, Lifelines Cohort
30 Study ⁷⁶, Giovanni Gambaro ¹⁰², Erik Ingelsson ^{103,104}, Lars Lind ¹⁰³, Paul IW de Bakker ^{105,106}, Mattijs E
31 Numans ^{107,106}, Ivan Brandslund ^{108,109}, Cramer Christensen ¹¹⁰, Eva RB Petersen ¹¹¹, Eeva Korpi-Hyövälti
32 ¹¹², Heikki Oksa ¹¹³, John C Chambers ^{48,49,114}, Jaspal S Kooner ^{49,115,114}, Alexandra IF Blakemore ^{42,43}, Steve
33 Franks ¹¹⁶, Marjo-Riitta Jarvelin ^{117,118,119,120}, Lise L Husemoen ¹²¹, Allan Linneberg ^{121,122,123}, Tea Skaaby
34 ¹²¹, Betina Thuesen ¹²¹, Fredrik Karpe ^{40,41}, Jaakko Tuomilehto ^{36,124,125,126}, Alex SF Doney ³⁹, Andrew D
35 Morris ¹²⁷, Colin NA Palmer ³⁹, Oddgeir Lingaas Holmen ^{128,129}, Kristian Hveem ^{128,130}, Cristen J Willer
36 ^{38,131,132}, Tiinamaija Tuomi ^{133,134}, Leif Groop ^{135,134}, AnneMari Käräjämäki ^{136,137}, Aarno Palotie ^{16,9,134,138},
37 Samuli Ripatti ^{134,139,30}, Veikko Salomaa ³⁶, Dewan S Alam ¹⁴⁰, Abdulla al Shafi Majumder ¹⁴¹, Emanuele Di
38 Angelantonio ^{1,54}, Rajiv Chowdhury ¹, Mark I McCarthy ^{40,41,20}, Neil Poulter ¹⁴², Alice V Stanton ¹⁴³, Peter
39 Sever ¹⁴², Philippe Amouyel ^{144,145,146,147}, Dominique Arveiler ¹⁴⁸, Stefan Blankenberg ^{149,150}, Jean Ferrières
40 ¹⁵¹, Frank Kee ¹⁵², Kari Kuulasmaa ³⁶, Martina Müller-Nurasyid ^{153,154,155}, Giovanni Veronesi ¹⁵⁶, Jarmo
41 Virtamo ³⁶, Panos Deloukas ^{21,157}, Wellcome Trust Case Control Consortium ⁷⁶, Paul Elliott ¹¹⁷,
42 Understanding Society Scientific Group ⁷⁶, Eleftheria Zeggini ³⁰, Sekar Kathiresan ^{56,158,159,9}, Olle Melander
43 ³³, Johanna Kuusisto ³², Markku Laakso ³², Sandosh Padmanabhan ⁹⁵, David Porteous ⁶⁶, Caroline Hayward
44 ³¹, Generation Scotland ¹⁶⁰, Francis S Collins ²⁹, Karen L Mohlke ¹⁶¹, Torben Hansen ¹¹, Oluf Pedersen ¹¹,
45 Michael Boehnke ^{13,12}, Heather M Stringham ^{13,12}, EPIC-CVD Consortium ⁷⁶, Philippe Frossard ²⁷,
46 Christopher Newton-Cheh ^{56,158}, CHARGE+ Exome Chip Blood Pressure Consortium ⁷⁶, Martin D Tobin ⁶,

47 Børge Grønne Nordestgaard ²⁶, T2D-GENES Consortium ⁷⁶, GoT2DGenes Consortium ⁷⁶, ExomeBP
48 Consortium ⁷⁶, CHD Exome+ Consortium ⁷⁶, Mark J Caulfield ^{4,5}, Anubha Mahajan ²⁰, Andrew P Morris
49 ^{20,7}, Maciej Tomaszewski ^{18,19,162}, Nilesh J Samani ^{18,19}, Danish Saleheen ^{28,27,1,167}, Folkert W Asselbergs
50 ^{15,101,163,167}, Cecilia M Lindgren ^{164,9,20,167}, John Danesh ^{1,165,54,167}, Louise V Wain ^{6,167}, Adam S Butterworth
51 ^{1,166,167}, Joanna MM Howson ^{1,167,168}, Patricia B Munroe ^{4,5,167,168}

52

- 53 1. Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of
54 Cambridge, Cambridge, UK
- 55 2. Medical Research Council Integrative Epidemiology Unit, School of Social and Community Medicine,
56 University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK
- 57 3. Centre for Cardiovascular Genetics, Institute of Cardiovascular Science, Rayne Building University
58 College London, London, UK
- 59 4. Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London,
60 UK
- 61 5. National Institute for Health Research Barts Cardiovascular Biomedical Research Unit, Queen Mary
62 University of London, London, UK
- 63 6. Department of Health Sciences, University of Leicester, Leicester, UK
- 64 7. Department of Biostatistics, University of Liverpool, Liverpool, UK
- 65 8. Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA
- 66 9. Program in Medical and Population Genetics, Broad Institute, 7 Cambridge Center, Cambridge,
67 Massachusetts, USA
- 68 10. Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, USA
- 69 11. The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical
70 Sciences, University of Copenhagen, Copenhagen, Denmark
- 71 12. Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA
- 72 13. Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA
- 73 14. Saw Swee Hock School of Public Health, National University of Singapore, National University Health
74 System, Singapore
- 75 15. Department of Cardiology, University Medical Center Utrecht, Utrecht, The Netherlands
- 76 16. Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital,
77 Boston, Massachusetts, USA
- 78 17. Genetics of Complex Traits, Institute of Biomedical and Clinical Science, University of Exeter Medical
79 School, Exeter, UK
- 80 18. Department of Cardiovascular Sciences, University of Leicester, Leicester, UK
- 81 19. National Institute for Health Research Leicester Biomedical Research Unit in Cardiovascular Disease,
82 Leicester, UK
- 83 20. Wellcome Trust Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford,
84 Oxford, UK
- 85 21. Heart Centre, William Harvey Research Institute, Barts and The London School of Medicine and
86 Dentistry, Queen Mary University of London, London, UK
- 87 22. Estonian Genome Center, University of Tartu, Tartu, Estonia
- 88 23. National Heart, Lung, and Blood Institute's and Boston University's Framingham Heart Study,
89 Framingham, Massachusetts, USA
- 90 24. Division of Statistical Genomics, Center for Genome Sciences and Systems Biology, Washington
91 University School of Medicine, St. Louis, Missouri, USA
- 92 25. Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA
- 93 26. Department of Clinical Biochemistry Herlev Hospital, Copenhagen University Hospital, Herlev,

94 Denmark
95 27. Centre for Non-Communicable Diseases, Karachi, Pakistan
96 28. Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of
97 Pennsylvania, Philadelphia, Pennsylvania, USA
98 29. Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, NIH,
99 Bethesda, Maryland, USA
100 30. Wellcome Trust Sanger Institute, Genome Campus, Hinxton, UK
101 31. Medical Research Council Human Genetics Unit, Medical Research Council Institute of Genetics and
102 Molecular Medicine, University of Edinburgh, Edinburgh, UK
103 32. Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio,
104 Finland
105 33. University of Lund, Department of Clinical Sciences, Malmö, Sweden
106 34. University of Verona, Department of Medicine, Verona, Italy
107 35. Department of Haematology, University of Cambridge, Cambridge, UK
108 36. Department of Health, National Institute for Health and Welfare, Helsinki, Finland
109 37. Institute of Molecular Medicine FIMM, University of Helsinki, Finland
110 38. Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann
111 Arbor, Michigan, USA
112 39. Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK
113 40. Oxford Centre for Diabetes, Endocrinology, and Metabolism, Radcliffe Department of Medicine,
114 University of Oxford, Oxford, UK
115 41. National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospital
116 Trusts, Oxford, UK
117 42. Section of Investigative Medicine, Imperial College London, London, UK
118 43. Department of Life Sciences, Brunel University London, London, UK
119 44. Institute of Biomedicine, Biocenter Oulu, University of Oulu, Oulu, Finland
120 45. Department of Gastroenterology and Metabolism, Poznan University of Medical Sciences, Poznan,
121 Poland
122 46. Hospital for Children and Adolescents, Helsinki University Central Hospital and University of Helsinki,
123 Helsinki, Finland
124 47. Department of Obstetrics and Gynaecology, Oulu University Hospital and University of Oulu, Oulu,
125 Finland
126 48. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London,
127 London, UK
128 49. Department of Cardiology, Ealing Hospital, Middlesex, UK
129 50. Medical Research Council Epidemiology Unit, University of Cambridge School of Clinical Medicine,
130 Box 285 Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, UK
131 51. Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden
132 52. Section of Biology and Genetics, Department of Neurosciences, Biomedicine and Movement Sciences,
133 University of Verona, Verona, Italy
134 53. Human Genetics, Wellcome Trust Sanger Institute, Hinxton, UK
135 54. The National Institute for Health Research Blood and Transplant Research Unit in Donor Health and
136 Genomics, University of Cambridge, Cambridge, UK
137 55. University Medical Center Groningen, University of Groningen, Department of Cardiology, The
138 Netherlands
139 56. Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA
140 57. Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece
141 58. Farr Institute of Health Informatics Research, Institute of Health Informatics, University College

142 London, London, UK

143 59. Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, Malmö,
144 Sweden

145 60. Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Science and
146 Technology, University of the Basque Country (UPV/EHU), Bilbao, Spain

147 61. University of Glasgow, Glasgow, UK

148 62. Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

149 63. Mr. De Craen suddenly passed away January 2016

150 64. Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands

151 65. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

152 66. Centre for Genomic and Experimental Medicine, Medical Research Council Institute of Genetics and
153 Molecular Medicine, University of Edinburgh, Edinburgh, UK

154 67. Department of Psychology, University of Edinburgh, Edinburgh, UK

155 68. Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia

156 69. Department of Twin Research and Genetic Epidemiology, King's College London, UK

157 70. Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University,
158 Athens, Greece

159 71. Department of Biobank Research, Umeå University, Umeå, Sweden

160 72. Kuopio Research Institute of Exercise Medicine, Kuopio, Finland

161 73. Section of Cardiology, Department of Medicine, Boston University Schools of Medicine and Public
162 Health, Boston, Massachusetts, USA

163 74. Sections of Preventive Medicine and Epidemiology, Department of Medicine, Boston University
164 Schools of Medicine and Public Health, Boston, Massachusetts, USA

165 75. Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The
166 Netherlands

167 76. A full list of members and affiliations appears in the Supplementary Note

168 77. Development Management and Planning, Pfizer Worldwide Research and Development

169 78. Pfizer Worldwide Research and Development, Stockholm, Sweden

170 79. Merck Research Laboratories, Genetics and Pharmacogenomics, Boston, Massachusetts, USA.

171 80. Merck Research Laboratories, Cardiometabolic Disease, Kenilworth, New Jersey, USA

172 81. CHDI Management/CHDI Foundation, Princeton, New Jersey, USA

173 82. Section of Computational Biomedicine, Department of Medicine, Boston University School of
174 Medicine, Boston, Massachusetts, USA

175 83. Institute of Epidemiology and Biobank Popgen, Kiel University, Kiel, Germany

176 84. Neurology Unit, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

177 85. University of North Carolina at Chapel Hill, Department of Epidemiology, Chapel Hill, North Carolina,
178 USA

179 86. Department of Public Health and Clinical Nutrition, University of Eastern Finland, Finland

180 87. Research Unit, Kuopio University Hospital, Kuopio, Finland

181 88. Institute of Biomedicine/Physiology, University of Eastern Finland, Kuopio Campus, Finland

182 89. Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland

183 90. Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics,
184 University of Edinburgh, Edinburgh, UK

185 91. Faculty of Medicine, University of Split, Croatia

186 92. Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden

187 93. Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA

188 94. Alzheimer Scotland Research Centre, University of Edinburgh, Edinburgh, UK

189 95. Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences,

190 University of Glasgow, Glasgow, UK
191 96. The Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands
192 97. Division of Endocrinology, Boston Children's Hospital, Boston, Massachusetts, USA
193 98. Institute of Molecular and Cell Biology, Tartu, Estonia
194 99. Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen,
195 The Netherlands
196 100. Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen,
197 The Netherlands
198 101. Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, The Netherlands
199 102. Division of Nephrology, Department of Internal Medicine and Medical Specialties, Columbus -
200 Gemelli University Hospital, Catholic University, Rome, Italy
201 103. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala
202 University, Uppsala, Sweden
203 104. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of
204 Medicine, Stanford, California, USA
205 105. Department of Medical Genetics, Center for Molecular Medicine, University Medical Center Utrecht,
206 Utrecht, The Netherlands
207 106. Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical
208 Center Utrecht, Utrecht, The Netherlands
209 107. Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The
210 Netherlands
211 108. Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark
212 109. Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark
213 110. Medical Department, Lillebaelt Hospital, Vejle, Denmark
214 111. Department of Clinical Immunology and Biochemistry, Lillebaelt Hospital, Vejle, Denmark
215 112. South Ostrobothnia Central Hospital, Seinäjoki, Finland
216 113. Tampere University Hospital, Tampere, Finland
217 114. Imperial College Healthcare NHS Trust, London, UK
218 115. National Heart and Lung Institute, Imperial College London, London, UK
219 116. Institute of Reproductive and Developmental Biology, Imperial College London, London, UK
220 117. Department of Epidemiology and Biostatistics, Medical Research Council Public Health England
221 Centre for Environment and Health, School of Public Health, Faculty of Medicine, Imperial College
222 London, St. Mary's Campus, London, UK
223 118. Centre for Life Course Epidemiology, Faculty of Medicine, University of Oulu, Oulu, Finland
224 119. Biocenter Oulu, University of Oulu, Oulu, Finland
225 120. Unit of Primary Care, Oulu University Hospital, Oulu, Finland
226 121. Research Centre for Prevention and Health, Capital Region of Denmark, Copenhagen, Denmark
227 122. Department of Clinical Experimental Research, Glostrup University Hospital, Glostrup, Denmark
228 123. Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen,
229 Copenhagen, Denmark
230 124. Dasman Diabetes Institute, Dasman, Kuwait
231 125. Centre for Vascular Prevention, Danube-University Krems, Krems, Austria
232 126. King Abdulaziz University, Jeddah, Saudi Arabia
233 127. School of Molecular, Genetic and Population Health Sciences, University of Edinburgh, Medical
234 School, Teviot Place, Edinburgh, UK
235 128. HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of
236 Science and Technology, Levanger, Norway
237 129. St. Olav Hospital, Trondheim University Hospital, Trondheim, Norway

238 130. Department of Medicine, Levanger Hospital, Nord- Trøndelag Health Trust, Levanger, Norway
239 131. Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor,
240 Michigan, USA
241 132. Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA
242 133. Folkhälsan Research Centre, Helsinki, Finland; Department of Endocrinology, Helsinki University
243 Central Hospital, Helsinki, Finland
244 134. Institute for Molecular Medicine Finland University of Helsinki, Helsinki, Finland
245 135. Department of Clinical Sciences, Diabetes and Endocrinology, Lund University Diabetes Centre,
246 Malmö, Sweden.
247 136. Department of Primary Health Care, Vaasa Central Hospital, Vaasa, Finland
248 137. Diabetes Center, Vaasa Health Care Center, Vaasa, Finland
249 138. Psychiatric and Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General
250 Hospital, Boston, Massachusetts, USA
251 139. Department of Public Health, University of Helsinki, Finland
252 140. ICDDR, B; Mohakhali, Dhaka, Bangladesh
253 141. National Institute of Cardiovascular Diseases, Sher-e-Bangla Nagar, Dhaka, Bangladesh
254 142. International Centre for Circulatory Health, Imperial College London, UK
255 143. Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Ireland
256 144. University of Lille, UMR1167, Risk Factors and Molecular Determinants of aging-related diseases,
257 Lille, France
258 145. Inserm, Lille, France
259 146. Centre Hospitalier Universitaire Lille, Public Health, Lille, France
260 147. Institut Pasteur de Lille, Lille, France
261 148. Department of Epidemiology and Public Health, EA 3430, University of Strasbourg, Strasbourg,
262 France
263 149. Department of General and Interventional Cardiology, University Heart Center Hamburg, Germany
264 150. University Medical Center Hamburg-Eppendorf, Hamburg, Germany
265 151. Department of Epidemiology, UMR 1027- INSERM, Toulouse University-CHU Toulouse, Toulouse,
266 France
267 152. Director, UKCRC Centre of Excellence for Public Health, Queens University, Belfast, Northern Ireland
268 153. Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for
269 Environmental Health, Neuherberg, Germany
270 154. Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität,
271 Munich, Germany
272 155. DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich,
273 Germany
274 156. Research Center in Epidemiology and Preventive Medicine, Department of Clinical and Experimental
275 Medicine, University of Insubria, Varese, Italy
276 157. Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-
277 HD), King Abdulaziz University, Jeddah, Saudi Arabia
278 158. Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA
279 159. Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA
280 160. A Collaboration between the University Medical Schools and NHS, Aberdeen, Dundee, Edinburgh and
281 Glasgow, UK
282 161. Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA
283 162. Institute of Cardiovascular Sciences, University of Manchester, Manchester, UK
284 163. Faculty of Population Health Sciences, Institute of Cardiovascular Science, University College London,
285 London, UK

286 164. The Big Data Institute at the Li Ka Shing Centre for Health Information and Discovery, University of
287 Oxford, Oxford, UK
288 165. Wellcome Trust Sanger Institute, Hinxton, UK
289 166. The National Institute for Health Research Blood and Transplant Research

290 167. Contributed equally to this work
291 168. Jointly supervised the work
292
293 Corresponding authors: Joanna M M Howson, jmmh2@medschl.cam.ac.uk and Patricia B Munroe,
294 p.b.munroe@qmul.ac.uk
295
296
297

Abstract

High blood pressure is a major risk factor for cardiovascular disease and premature death. However, there is limited knowledge on specific causal genes and pathways. To better understand the genetics of blood pressure, we genotyped 242,296 rare, low-frequency and common genetic variants in up to ~192,000 individuals, and used ~155,063 samples for independent replication. We identified 31 novel blood pressure or hypertension associated genetic regions in the general population, including three rare missense variants in *RBM47*, *COL21A1* and *RRAS* with larger effects ($>1.5\text{mmHg/allele}$) than common variants. Multiple rare, nonsense and missense variant associations were found in *A2ML1* and a low-frequency nonsense variant in *ENPEP* was identified. Our data extend the spectrum of allelic variation underlying blood pressure traits and hypertension, provide new insights into the pathophysiology of hypertension and indicate new targets for clinical intervention.

310 **Introduction**

311 High blood pressure (BP) or hypertension is a highly prevalent chronic disorder. It is estimated to be
312 responsible for a larger proportion of global disease burden and premature mortality than any other disease
313 risk factor¹. Elevated systolic and/or diastolic BP increases the risk of several cardiovascular disorders
314 including stroke, coronary heart disease (CHD), heart failure, peripheral arterial disease and abdominal
315 aortic aneurysms². BP is a complex, heritable, polygenic phenotype for which genome-wide association
316 studies (GWAS) have identified over 67 genetic regions associated with BP and/or hypertension to date³⁻¹¹.
317 These variants are common (minor allele frequency, $MAF \geq 0.05$), mostly map to intronic or intergenic
318 regions, with the causal alleles and genes not readily identified due to linkage disequilibrium (LD)^{4,5}, and
319 explain only ~2% of trait variance¹². Low-frequency ($0.01 < MAF < 0.05$) and rare ($MAF \leq 0.01$) single
320 nucleotide variants (SNVs), predominantly unexplored by GWAS may have larger phenotypic effects than
321 common SNVs¹³, and may help to explain the missing heritability, and identify causative genes as
322 demonstrated previously¹⁴.

323 To identify novel coding variants and loci influencing BP traits and hypertension we performed the largest
324 meta-analysis to date that included a total of ~350,000 individuals, directly genotyped with the Exome chip.
325 The Exome chip contains ~240,000 mostly rare and low-frequency variants (Methods). A single-variant
326 discovery analysis was performed, and candidate SNVs were taken forward for validation using independent
327 replication samples. Gene-based tests were used to identify BP associated genes harboring multiple rare
328 variant associations. We next assessed whether the newly identified BP associated SNVs were associated
329 with expression levels of nearby genes, and tested these variants in aggregate for a causal association of BP
330 with other cardiovascular traits and risk factors. Our findings highlight the contribution of rare variants in
331 the aetiology of blood pressure in the general population, and provide new insights into the pathophysiology
332 of hypertension.

333

334 **Results**

335 **Discovery of single variant BP associations**

336 We genotyped 192,763 individuals from 51 studies, and assessed association of 242,296 SNVs with
337 diastolic BP (DBP), systolic BP (SBP), pulse pressure (PP) and hypertension (HTN; Supplementary Tables
338 1, 2 and 3; Methods). An overview of the SNV discovery study design is given in Figure 1. A fixed effects
339 meta-analysis for each trait was performed using study-level association summary statistics from i) samples
340 of European (EUR) ancestry (up to 165,276 individuals), and ii) a trans-ethnic meta-analysis of the EUR and
341 additional South Asian (SAS) ancestry samples (EUR_SAS; up to 192,763 individuals). Two analyses of
342 DBP, SBP and PP were performed, one in which the trait was inverse normal transformed and a second in
343 which the raw phenotype was analysed. Both sets of results were consistent (Methods), therefore to
344 minimise sensitivity to deviations from normality in the analysis of rare variants, the results from the
345 analyses of the transformed traits were used for discovery. Strong correlations between the BP traits were
346 observed across studies (Methods), hence no adjustment of significance thresholds for independent trait
347 testing was applied.

348 The discovery meta-analyses identified 50 genomic regions with genome-wide significant (GWS) evidence
349 of association with at least one of the four BP traits tested ($P < 5 \times 10^{-8}$; Supplementary Table 4). There were
350 45 regions associated in the EUR_SAS samples, of which 13 were novel (Figure 2). An additional five
351 regions were GWS in the EUR only meta-analyses of which two were novel (Supplementary Figure 1). In
352 total, 16 genomic regions were identified that were GWS for at least one BP trait that have not been
353 previously reported.

354 **Replication of single variant BP associations**

355 Next we sought support for our findings, in an independent replication dataset comprising of 18 studies, 16
356 of which were from the Cohorts for Heart and Aging Research in Genomic Epidemiology+ (CHARGE+)
357 exome chip blood pressure consortium (Figure 1; Liu *et al.* Nature Genetics, *submitted*). Variants were
358 selected for replication first using the larger (transformed) EUR_SAS data, with additional variants from the
359 (transformed) EUR data also selected. SNVs were selected if they mapped outside of known BP genomic
360 regions and had $MAF \geq 0.05$ and $P < 1 \times 10^{-5}$ or $MAF < 0.05$ and $P < 1 \times 10^{-4}$ with at least one BP trait, *i.e.*
361 choosing a lower significance threshold for the selection of rare variants (full details of the selection criteria
362 are provided in the Methods). In total 81 candidate SNVs were selected for replication (Supplementary

Table 5). Eighty variants were selected from EUR_SAS (transformed) results and one SNV at the *ZNF101* locus from the EUR (transformed) analyses. The results for EUR_SAS and EUR were consistent (association statistics were correlated, $\rho=0.9$ across ancestries for each of the traits). Of the 81 variants, 30 SNVs were selected for association with DBP as the primary trait, 26 for SBP, 19 for PP and 6 for HTN, with the primary trait defined as the BP trait with the smallest association *P*-value in the EUR-SAS discovery analyses.

Meta-analyses were performed on results from analyses of untransformed DBP, SBP, PP and HTN (as only results of untransformed traits were available from CHARGE+) in (i) up to 125,713 individuals of EUR descent, and (ii) up to 155,063 individuals of multiple ethnicities (4,632 of Hispanic descent, 22,077 of African American descent, 2,641 SAS samples with the remainder EUR; Figure 1). Given that a large proportion of the ancestries in the trans-ethnic meta-analyses were not included in our discovery samples, we used the EUR meta-analyses as the main data set for replication, but we also report any additional associations identified within the larger trans-ethnic dataset.

Novel BP-SNV associations were identified based on two criteria (Figure 1; Methods). Firstly, replication of the primary BP trait-SNV association was sought at a Bonferroni adjusted *P*-value threshold in the replication data ($P \leq 6.17 \times 10^{-4}$, assuming $\alpha=0.05$ for 81 SNVs tested and same direction of effect; Methods) without the need for GWS. Secondly, meta-analyses of discovery and replication results across all four (untransformed) BP traits were performed to assess the overall level of support across all samples for the 81 candidate SNVs; those BP-SNV associations that were GWS (with statistical support in the replication studies; $P < 0.05$ and the same direction of effect) were also declared as novel.

Seventeen SNV-BP associations formally replicated with concordant direction of effect at a Bonferroni adjusted significance level for the primary trait. Fourteen were in the EUR meta-analyses, and amongst these was a rare non-synonymous (ns) SNV mapping to *COL21A1* (Table 1, Supplementary Table 6). Three associations were in the trans-ethnic meta-analyses, these included two rare nsSNVs in *RBM47* and *RRAS* (Table 1, Supplementary Table 7; Methods).

389

In addition to the 17 SNV-BP trait associations that formally replicated, we identified 13 further SNV-associations that were GWS in the combined (discovery and replication) meta-analyses. Ten of these were GWS in the combined EUR analyses, (Table 2; Supplementary Tables 6 and 8a), and three were GWS in the combined trans-ethnic meta-analyses (Table 2; Supplementary Tables 7 and 8b).

394

This gives a total of 30 novel SNV-BP associations (15 SNV-DBP, 9 SNV-SBP and 6 SNV-PP; Tables 1 and 2; Supplementary Figures 2 and 3). Five of the SNVs were GWS with more than one BP trait (Figure 3: Tables 1 and 2; Supplementary Table 8). Four loci (*CERS5*, *TBX2*, *RGL3* and *OBFC1*) had GWS associations with HTN in addition to GWS associations with DBP and SBP. The *PRKAG1* locus had GWS associations with both SBP and PP.

400

Conditional analyses were performed to identify secondary signals of association within the novel BP loci. The RAREMETALWORKER (RMW) package (Methods)¹⁵ allows conditional analyses to be performed using summary level data. Hence, analyses of the transformed primary traits and HTN were re-run in RMW across the discovery studies (Figure 4). The results of the RMW single variant tests were consistent with the initial discovery analyses (Supplementary Information). Given the RMW analyses were based on our discovery samples, the larger EUR-SAS data was used as the main analysis to increase power, but we also report any additional associations with evidence in EUR.

408

We identified secondary independent signals of association in four loci, *PREX1*, *PRKAG1* and *RRP1B* within the EUR_SAS analyses and *COL21A1* in the EUR analyses ($P_{\text{conditional}} < 1 \times 10^{-4}$, Bonferroni adjusted for ~500 variants within each region; Methods; Supplementary Tables 9 and 10). Three independent association signals were identified in the *MYH6* locus in the EUR_SAS analyses (Supplementary Table 11).

413

414 **Gene-based BP associations**

To improve statistical power to detect associations in genes harbouring rare variants, analytical methods that combine effects of variants across a gene into a single test have been devised and are implemented in the

416

RMW package¹⁵. We applied the gene-based sequence kernel association test (SKAT)¹⁶ and Burden tests¹⁷ to the RMW dataset (MAF<0.05 or MAF<0.01; Figure 4; Methods). One previously unidentified BP gene (*A2ML1*) was associated with HTN ($P=7.73 \times 10^{-7}$) in the EUR_SAS studies and also in EUR studies (Supplementary Table 12; Bonferroni-corrected threshold of significance $P<2.8 \times 10^{-6}$, after adjusting for 17,996 genes tested, Methods). The gene showed residual association with the primary BP trait after conditioning on the most associated SNV in the gene ($P_{\text{conditional}}=5.00 \times 10^{-4}$; Supplementary Table 12), suggesting that the association is due to multiple rare variants in the gene. One nonsense variant (rs199651558, p.Arg893*, MAF= 3.5×10^{-4}) was observed, and there were multiple missense variants (Figure 5). *A2ML1* encodes alpha-2-macroglobulin-like 1 protein, and is a member of the alpha macroglobulin superfamily, which comprises protease inhibitors targeting a wide range of substrates. Mutations in this gene are associated with a disorder clinically related to Noonan syndrome, a developmental disorder which involves cardiac abnormalities¹⁸. We sought replication in the CHARGE+ studies for this gene, however there was no evidence of association with HTN ($P=0.45$). Given the very low frequencies of the variants involved, however, studies in which the variants are polymorphic will be required to replicate the association with HTN. The DBH gene was found to be associated with DBP using the SKAT test ($P=2.88 \times 10^{-6}$). However, this was not due to multiple rare variants as the association was driven by rs77273740 (Supplementary Table 5) and the SNV was not validated in the replication samples.

Rare and common variant associations in established BP loci

Of the 67 established BP loci, 35 loci were on the Exome chip (N=43 SNVs or close proxies $r^2>0.7$). All 43 SNVs had at least nominal evidence of association with BP in our discovery samples ($P<0.01$; Supplementary Table 13). We also assessed if any of the established BP loci contained coding variants that are associated with BP traits and in LD ($r^2>0.2$) with the known BP variants on the Exome chip (Supplementary Table 13), using the 1000G phase 3 release for LD annotation. Focusing on SNVs that were GWS for any BP trait from our transformed discovery data for either ancestry, there were 25 coding variants, of which 6 were predicted to be damaging at loci labelled *CDC25A*, *SLC39A8*, *HFE*, *ULK4*, *ST7L-CAPZA1-MOV10* and *CYP11A1-ULK3*. Three of these are published variants at loci labelled *SLC39A8*, *HFE* and *ST7-CAPZA1-MOV10*. At *CYP11A1-ULK3*, the coding variant was in moderate LD with the reported

variant, but was less significantly associated with DBP in our EUR_SAS dataset ($P=2.24 \times 10^{-8}$ compared to $P=1.68 \times 10^{-15}$ for the published variant). At the *ULK4* locus the predicted damaging coding variant had similar association as the published coding variant (predicted to be benign), and prior work has already indicated several associated nsSNVs in strong LD in *ULK4*¹⁹. The nsSNV within the *CDC25A* locus (rs11718350 in *SPINK8*) had similar association with DBP as the intergenic published SNV in our EUR_SAS dataset ($P=2.00 \times 10^{-8}$ compared to $P=2.27 \times 10^{-8}$ for the published variant). Overall at least 5 of the known loci are consistent with having a coding causal variant.

Gene-based SKAT tests of all genes that map within 1 Mb of a previously reported SNV association (Supplementary Table 14), indicated no genes with multiple rare or low-frequency variant associations. Single variant conditional analyses showed that rs33966350, a rare nonsense variant in *ENPEP* (MAF=0.01) was associated with SBP ($P_{\text{conditional}}=1.61 \times 10^{-5}$) in the EUR_SAS samples (Supplementary Tables 14 and 15; Methods) independently of the known SNV (rs6825911). *ENPEP* encodes aminopeptidase A (APA) an enzyme of the renin-angiotensin-aldosterone system (RAAS) that converts angiotensin II (AngII) to AngIII. There were no other established loci with convincing low-frequency or rare SNV associations in the EUR_SAS samples. However, *HOXC4*, had evidence of a second independent signal with a rare missense SNV in EUR samples (rs78731604; MAF=0.005, $P_{\text{conditional}}=5.76 \times 10^{-5}$; Supplementary Table 15). The secondary signal in the *HOXC4* region, mapped to *CALCOCO1*, ~300kb from the known SNV. The gene association (MAF \leq 0.01, $P=2.37 \times 10^{-5}$) was below the required significance threshold and attributable to rs78731604, which is not predicted to have detrimental effects on protein structure. Therefore, replication of this association is required. Three loci (*ST7L-CAPZA1-MOV10*, *FIGN-GRB14*, and *TBX5-TBX3*) had evidence of a second independent signal in the region in EUR_SAS samples with a common variant ($P_{\text{conditional}} < 1 \times 10^{-4}$; Supplementary Table 15) that has not been previously reported.

Having identified 30 novel loci associated with BP traits, as well as additional new independent SNVs at four novel loci and five known loci, we calculated the percent of the trait variance explained (Methods). This was 2.08%/2.11%/1.15% for SBP/DBP/PP for the 43 previously reported BP-SNVs covered in our dataset, increasing to 3.38%/3.41%/2.08% respectively with the inclusion of the 30 lead SNVs from novel loci, plus new independent SNV-BP associations identified from novel and known loci.

473 **Effect of BP SNVs on cardiovascular traits & risk factors**

474 Amongst our novel BP-SNV associations, some have previously been reported to be associated with other
 475 cardiovascular traits and risk factors (Supplementary Table 16); these include coronary heart disease (CHD:
 476 *PHACTR1*, *ABO*)^{20,21}, QT interval (*RNF207*)²², heart rate (*MYH6*)²³, and cholesterol levels (2q36.3, *ABO*,
 477 *ZNF101*)²⁴.

478 To test the impact of BP variants on cardiovascular endpoints and risk factors we created three weighted
 479 genetic risk scores (GRS) according to SBP/DBP/PP based on the newly identified and previously published
 480 BP variants (up to N=125; Methods). The GRS models were used to test the causal effect of BP on the
 481 following traits: ischemic stroke (including the subtypes, cardiometabolic, large and small vessel²⁵), CHD,
 482 heart failure,²⁶ left ventricular mass²⁷, left ventricular wall thickness²⁷, high-density lipoprotein cholesterol
 483 (HDL-c), low-density lipoprotein (LDL-c), triglycerides, total cholesterol, body mass index (BMI), waist-
 484 hip ratio adjusted BMI, height and estimated glomerular filtration rate (eGFR) (Methods). As expected, BP
 485 was positively associated with increased CHD risk (OR [95% CI]=1.39[1.22-1.59] per 10mmHg increase in
 486 SBP, $P=6.07 \times 10^{-7}$; 1.62[1.28-2.05] per 10mmHg increase in DBP, $P=5.99 \times 10^{-5}$; 1.70[1.34-2.16] per
 487 10mmHg increase in PP, $P=1.20 \times 10^{-5}$; Table 3), and increased risk of ischemic stroke (OR [95%
 488 CI]=1.93[1.47-2.55] per 10mmHg increase in DBP, $P=2.81 \times 10^{-6}$; 1.57[1.35-1.84] per 10mmHg increase in
 489 SBP, $P=1.16 \times 10^{-8}$; 2.12[1.58-2.84] per 10mmHg increase in PP, $P=5.35 \times 10^{-7}$). The positive association with
 490 ischemic stroke was primarily due to large vessel stroke (Table 3). DBP and SBP were also positively
 491 associated with left ventricular mass (9.57 [3.98-15.17] gram increase per 10mmHg increase in DBP,
 492 $P=8.02 \times 10^{-4}$ and 5.13 [1.77-8.48] gram increase per 10mmHg increase in SBP, $P=0.0027$) and left
 493 ventricular wall thickness (0.10 [0.06-0.13] cm increase per 10mmHg increase in DBP, $P=1.88 \times 10^{-8}$ and
 494 0.05 [0.03-0.07] cm increase per 10mmHg increase in SBP, $P=5.52 \times 10^{-6}$, Table 3). There was no convincing
 495 evidence to support the BP associated variants having an effect on lipid levels ($P>0.1$), BMI ($P>0.005$),
 496 waist hip ratio adjusted BMI ($P>0.1$), height ($P>0.06$), eGFR ($P>0.02$) or heart failure ($P>0.04$). The causal
 497 associations with CHD, stroke, and left ventricular measures augment the results from a previous association
 498 analysis using 29 BP variants²⁸. Our data strongly support the previous observations of no causal

relationship between BP and eGFR. Lack of evidence of a BP effect with heart failure may only be due to lack of power, as the association was in the expected direction.

Possible functional variants at BP loci and candidate genes

Twenty-six of our newly discovered BP associated SNVs had $MAF \geq 0.05$ and therefore due to extensive LD with other SNVs not genotyped on the Exome array, identifying the causal genes requires additional information. If a SNV is associated with increased or decreased expression of a particular gene, *i.e.* it is an expression quantitative trait locus (eQTL) this suggests the gene on which the SNV acts could be in the causal pathway. To help identify potential candidate causal genes in the novel BP loci (Supplementary Table 9), information from publicly available eQTL databases was investigated (MuTHER for LCL, adipose and skin and GTEx for nine tissues including the heart and tibial artery; Methods).

The DBP increasing allele of the nsSNV, rs7302981-A, was associated with increased expression of *CERS5* in: LCLs ($P_{MuTHER}=3.13 \times 10^{-72}$) skin ($P_{MuTHER}=2.40 \times 10^{-58}$) adipose ($P_{MuTHER}=2.87 \times 10^{-54}$) and nerve ($P_{GTEx}=4.5 \times 10^{-12}$) (Supplementary Figure 4). Additional testing (Methods) provided no evidence against colocalisation of the eQTL and DBP association signals, implicating *CERS5* as a candidate causal gene for this DBP locus. *CERS5* (LAG1 homolog, ceramide synthase 5) is involved in the synthesis of ceramide, a lipid molecule involved in several cellular signaling pathways. *Cers5* knockdown has been shown to reduce cardiomyocyte hypertrophy in mouse models²⁹. However, it is unclear whether the blood pressure raising effects at this locus are the cause or result of any potential effects on cardiac hypertrophy. Future studies investigating this locus in relation to parameters of cardiac hypertrophy and function (*e.g.* ventricular wall thickness) should help address this question.

The DBP raising allele of the nsSNV (rs867186-A) was associated with increased expression of *PROCR* in adipose tissue ($P_{MuTHER}=3.24 \times 10^{-15}$) and skin ($P_{MuTHER}=1.01 \times 10^{-11}$) (Supplementary Figure 4). There was no evidence against colocalisation of the eQTL and DBP association thus supporting *PROCR* as a candidate causal gene. *PROCR* encodes the Endothelial Protein C receptor, a serine protease involved in the blood

coagulation pathway, and rs867186 has previously been associated with coagulation and haematological factors.^{30,31} The PP decreasing allele of, rs10407022-T, which is predicted to have detrimental effects on protein structure (Methods) was associated with increased expression of *AMH* in muscle ($P_{\text{GTEx}}=9.95 \times 10^{-15}$), thyroid ($P_{\text{GTEx}}=8.54 \times 10^{-7}$), nerve ($P_{\text{GTEx}}=7.15 \times 10^{-8}$), tibial artery ($P_{\text{GTEx}}=6.46 \times 10^{-9}$), adipose ($P_{\text{GTEx}}=4.69 \times 10^{-7}$), and skin ($P_{\text{GTEx}}=5.88 \times 10^{-8}$) (Supplementary Figure 4). There was no evidence against colocalisation of the eQTL and PP association, which supports *AMH* as a candidate causal gene for PP. Low AMH levels have been previously associated with hypertensive status in women with the protein acting as a marker of ovarian reserve³². The intergenic SBP raising allele of rs4728142-A was associated with reduced expression of *IRF5* in skin ($P_{\text{MuTHER}}=5.24 \times 10^{-31}$) and LCLs ($P_{\text{MuTHER}}=1.39 \times 10^{-34}$), whole blood ($P_{\text{GTEx}}=3.12 \times 10^{-7}$) and tibial artery ($P_{\text{GTEx}}=1.71 \times 10^{-7}$).

Three novel rare nsSNVs were identified that map to *RBM47*, *RRAS* (both associated with SBP) and *COL21A1* (associated with PP). They had larger effect sizes than common variant associations (>1.5 mmHg per allele; Supplementary Figure 5) and were predicted to have detrimental effects on protein structure (Supplementary Table 16; Methods). In *RBM47*, rs35529250 (p.Gly538Arg) is located in a highly conserved region of the gene and was most strongly associated with SBP (MAF=0.008; +1.59 mmHg per T allele; $P=5.90 \times 10^{-9}$). *RBM47* encodes the RNA binding motif protein 47 and is responsible for post-transcriptional regulation of RNA, through its direct and selective binding with the molecule.³³ In *RRAS*, rs61760904 (p.Asp133Asn) was most strongly associated with SBP (MAF=0.007; +1.51 mmHg per T allele; $P=8.45 \times 10^{-8}$). *RRAS* encodes a small GTPase belonging to the Ras subfamily of proteins H-RAS, N-RAS, and K-RAS and has been implicated in actin cytoskeleton remodelling, and controlling cell proliferation, migration and cycle processes³⁴. The nsSNV in *COL21A1* (rs200999181, p.Gly665Val) was most strongly associated with PP (MAF=0.001; +3.14 mmHg per A allele; $P=1.93 \times 10^{-9}$). *COL21A1* encodes the collagen alpha-1 chain precursor of type XXI collagen, a member of the FACIT (fibril-associated collagens with an interrupted triple helix) family of proteins³⁵. The gene is detected in many tissues, including the heart and aorta. Based on our results, these three genes represent good candidates for functional follow-up. However, due to the incomplete coverage of all SNVs across the region on the Exome chip, it is possible that other non-genotyped SNVs may better explain some of these associations. We therefore checked for variants in LD

553 ($r^2 > 0.3$) with these three rare nsSNVs in the UK10K + 1000G dataset³⁶ to ascertain if there are other
554 candidate SNVs at these loci (Supplementary Table 17). There were no SNVs within 1Mb of the *RBM47*
555 locus in LD with the BP associated SNV. At the *COL21A1* locus there were only SNVs in moderate LD, and
556 these were annotated as intronic, intergenic or in the 5'UTR. At the *RRAS* locus, there were two SNVs in
557 strong LD with the BP associated SNV, which both mapped to introns of *SCAF1* and are not predicted to be
558 damaging. All SNVs in LD at both loci were rare as expected (Supplementary Table 17) supporting a role
559 for rare variants. Hence, the rare BP associated nsSNVs at *RBM47*, *COL21A1* and *RRAS* remain the best
560 causal candidates.

561

562 **Pathway and network analyses**

563 To identify connected gene sets and pathways implicated by the BP associated genes we used Meta-Analysis
564 Gene-set Enrichment of variant Associations (MAGENTA)³⁷ and GeneGo MetaCore (Thomson Reuters,
565 UK). MAGENTA tests for over-representation of BP associated genes in pre-annotated pathways (gene
566 sets) (Methods and Supplementary Table 18a). GeneGo Metacore identifies potential gene networks. The
567 MAGENTA analysis was used for hypothesis generation and results were compared with the GeneGo
568 Metacore outputs to cross-validate findings.

569 Using MAGENTA there was an enrichment ($P < 0.01$ and $FDR < 5\%$ in either the EUR_SAS or the EUR
570 participants) of six gene sets with DBP, three gene sets with HTN and two gene sets for SBP
571 (Supplementary Table 18b). The RNA polymerase I promoter clearance (chromatin modification) pathway
572 showed the most evidence of enrichment with genes associated with DBP ($P_{\text{Reactome}} = 8.4 \times 10^{-5}$, $FDR = 2.48\%$).
573 NOTCH signalling was the most associated pathway with SBP ($P_{\text{Reactome}} = 3.00 \times 10^{-4}$, $FDR = 5\%$) driven by
574 associations at the *FURIN* gene. The inorganic cation anion solute carrier (SLC) transporter pathway had
575 the most evidence of enrichment by HTN associated genes ($P_{\text{Reactome}} = 8.00 \times 10^{-6}$, $FDR = 2.13\%$).

576 Using GeneGo MetaCore, five network processes were enriched ($FDR < 5\%$; Methods; Supplementary
577 Tables 19 and 20). These included several networks with genes known to influence vascular tone and BP:
578 inflammation signalling, $P = 1.14 \times 10^{-4}$ and blood vessel development $P = 2.34 \times 10^{-4}$. The transcription and
579 chromatin modification network ($P = 2.85 \times 10^{-4}$) was also enriched, a pathway that was also highlighted in the

MAGENTA analysis, with overlap of the same histone genes (*HIST1H4C*, *HIST1H2AC*, *HIST1H2BC*, *HIST1H1T*) and has also been recently reported in an integrative network analysis of published BP loci and whole blood expression profiling³⁸. Two cardiac development pathways were enriched: the oxidative stress-driven (ROS/NADPH) ($P=4.12 \times 10^{-4}$) and the Wnt/ β -catenin/integrin-driven ($P=0.0010$). Both these cardiac development pathways include the *MYH6*, *MYH7*, and *TBX2* genes, revealing a potential overlap with cardiomyopathies and hypertension, and suggesting some similarity in the underlying biological mechanisms.

587

588 Discussion

By conducting the largest ever genetic study of BP, we identified further novel common variants with small effects on BP traits, similar to what has been observed for obesity and height^{39,40}. More importantly, our study identified some of the first rare coding variants of strong effect ($>1.5\text{mmHg}$) that are robustly associated with BP traits in the general population, complementing and extending the previous discovery and characterisation of variants underlying rare Mendelian disorders of blood pressure regulation⁴¹. Using SNV associations in 17 genes reported to be associated with monogenic disorders of blood pressure (Methods) we found no convincing evidence of enrichment ($P_{\text{enrichment}}=0.044$). This suggests that BP control in the general population may occur through different pathways to monogenic disorders of BP re-enforcing the importance of our study findings. The identification of 30 novel BP loci plus further new independent secondary signals within four novel and five known loci (Methods) has augmented the trait variance explained by 1.3%, 1.2% and 0.93% for SBP, DBP and PP respectively within our data-set. This suggests that with substantially larger sample sizes, for example through UK BioBank⁴², we expect to identify 1000s more loci associated with BP traits, and replicate more of our discovery SNV associations that are not yet validated in the current report.

The discovery of rare missense variants has implicated several interesting candidate genes, which are often difficult to identify from common variant GWAS, and should therefore lead to more rapidly actionable biology. *A2ML1*, *COL21A1*, *RRAS* and *RBM47* all warrant further follow-up studies to define the role of

606 these genes in regulation of BP traits, as well as functional studies to understand their mechanisms of action.
607 *COL21A1* and *RRAS* warrant particular interest since both are involved in blood vessel remodelling, a
608 pathway of known aetiological relevance to hypertension.

609 We observed a rare nonsense SBP associated variant in *ENPEP* (rs33966350; p.Trp317*): this overlaps a
610 highly conserved region of both the gene and protein and is predicted to result in either a truncated protein
611 with reduced catalytic function or is subject to nonsense mediated RNA decay. ENPEP converts angiotensin
612 II (AngII) to Ang-III. AngII activates the angiotensin 1 (AT1) receptor resulting in vasoconstriction, while
613 AngIII activates the angiotensin 2 (AT2) receptor that promotes vasodilation and protects against
614 hypertension.⁴³ The predicted truncated protein may lead to predominant AngII signaling in the body, and
615 increases in BP. This new observation could potentially inform therapeutic strategies. Of note, angiotensin-
616 converting-enzyme (ACE) inhibitors are commonly used in the treatment of hypertension. However, patients
617 who suffer from adverse reactions to ACE inhibitors, such as dry cough and skin rash, would benefit from
618 alternative drugs that target RAAS. Murine studies have shown that in the brain, AngIII is the preferred AT1
619 agonist that promotes vasoconstriction and increases blood pressure, as opposed to AngII in the peripheral
620 system. These results have motivated the development of brain specific APA inhibitors to treat
621 hypertension⁴⁴. Our results confirm APAs, such as ENPEP, as a valid target to modify blood pressure, but
622 suggest that long-term systemic reduction in APA activity may lead to an increase in blood pressure. Future
623 studies are needed to examine the effects of the p.Trp317* variant on the RAAS system, specifically in the
624 brain and peripheral vasculature, in order to test the benefits of the proposed therapeutic strategy in humans.

625 In addition to highlighting new genes in pathways of established relevance to BP and hypertension, and
626 identifying new pathways, we have also identified multiple signals at new loci. For example, there are three
627 distinct signals at the locus containing the *MYH6/MYH7* genes, and we note that *TBX2* maps to one of the
628 novel regions. These genes are related to cardiac development and/or cardiomyopathies, and provide an
629 insight into the shared inheritance of multiple complex traits. Unravelling the causal networks within these
630 polygenic pathways may provide opportunities for novel therapies to treat or prevent both hypertension and
631 cardiomyopathies.

633 **URLs**

634 Exome chip design information: http://genome.sph.umich.edu/wiki/Exome_Chip_Design

635 [RareMetalWorker](http://genome.sph.umich.edu/wiki/RAREMETALWORKER) information: <http://genome.sph.umich.edu/wiki/RAREMETALWORKER>

636 [Summary SNV association results: http://www.phenoscanter.medschl.cam.ac.uk](http://www.phenoscanter.medschl.cam.ac.uk)

637 [Databases used for variant annotation: http://www.ncbi.nlm.nih.gov/SNP/](http://www.ncbi.nlm.nih.gov/SNP/)

638 <http://www.ensembl.org/info/docs/tools/index.html> and <http://evs.gs.washington.edu/EVS/>

639 UCSC reference file used for annotation of variants with gene and exon information:

640 <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/refFlat.txt.gz>

641 Databases used for pathway analysis: MAGENTA (<https://www.broadinstitute.org/mpg/magenta/>) and

642 THOMSON REUTERS MetaCore™ Single Experiment Analysis workflow tool

643 ([http://thomsonreuters.com/en/products-services/pharma-life-sciences/pharmaceutical-](http://thomsonreuters.com/en/products-services/pharma-life-sciences/pharmaceutical-research/metacore.html)

644 [research/metacore.html](http://thomsonreuters.com/en/products-services/pharma-life-sciences/pharmaceutical-research/metacore.html)).

645

646 **Acknowledgements**

647 **1. CHD Exome+ Consortium**

648 **CCHS, CGPS, CIHDS:** We thank participants and staff of the Copenhagen City Heart Study, Copenhagen

649 Ischemic Heart Disease Study, and the Copenhagen General Population Study for their important

650 contributions.

651 **EPIC-InterAct:** Funding for the InterAct project was provided by the EU FP6 programme (grant number

652 LSHM_CT_2006_037197). We thank all EPIC participants and staff for their contribution to the study. We

653 thank the lab team at the MRC Epidemiology Unit for sample management and Nicola Kerrison for data

654 management.

655 **EPIC-CVD:** CHD case ascertainment and validation, genotyping, and clinical chemistry assays in EPIC-

656 CVD were principally supported by grants awarded to the University of Cambridge from the EU Framework

657 Programme 7 (HEALTH-F2-2012-279233), the UK Medical Research Council (G0800270) and British Heart
658 Foundation (SP/09/002), and the European Research Council (268834). We thank all EPIC participants and
659 staff for their contribution to the study, the laboratory teams at the Medical Research Council Epidemiology
660 Unit for sample management and Cambridge Genomic Services for genotyping, Sarah Spackman for data
661 management, and the team at the EPIC-CVD Coordinating Centre for study coordination and administration.

662 **MORGAM:** This work has been sustained by the MORGAM Project's recent funding: European Union FP
663 7 projects ENGAGE (HEALTH-F4-2007-201413), CHANCES (HEALTH-F3-2010-242244) and
664 BiomarCaRE (278913). This has supported central coordination, workshops and part of the activities of the
665 The MORGAM Data Centre, at THL in Helsinki, Finland. The MORGAM Participating Centres are funded
666 by regional and national governments, research councils, charities, and other local sources.

667 **WOSCOPS/PROSPER:** The research leading to these results has received funding from the European
668 Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° HEALTH-F2-2009-
669 223004

670 **BRAVE:** The BRAVE study genetic epidemiology working group is a collaboration between the
671 Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge,
672 UK, the Centre for Control of Chronic Diseases, icddr,b, Dhaka, Bangladesh and the National Institute of
673 Cardiovascular Diseases, Dhaka, Bangladesh.

674 **PROMIS:** We are thankful to all the study participants in Pakistan. Recruitment in PROMIS was funded
675 through grants available to investigators at the Center for Non-Communicable Diseases, Pakistan (Danish
676 Saleheen and Philippe Frossard) and investigators at the University of Cambridge, UK (Danish Saleheen and
677 John Danesh). Field-work, genotyping, and standard clinical chemistry assays in PROMIS were principally
678 supported by grants awarded to the University of Cambridge from the British Heart Foundation, UK Medical
679 Research Council, Wellcome Trust, EU Framework 6-funded Bloodomics Integrated Project, Pfizer, Novartis,
680 and Merck. We would like to acknowledge the contributions made by the following individuals who were
681 involved in the field work and other administrative aspects of the study: Mohammad Zeeshan Ozair, Usman
682 Ahmed, Abdul Hakeem, Hamza Khalid, Kamran Shahid, Fahad Shuja, Ali Kazmi, Mustafa Qadir Hameed,

Naeem Khan, Sadiq Khan, Ayaz Ali, Madad Ali, Saeed Ahmed, Muhammad Waqar Khan, Muhammad Razaq Khan, Abdul Ghafoor, Mir Alam, Riazuddin, Muhammad Irshad Javed, Abdul Ghaffar, Tanveer Baig Mirza, Muhammad Shahid, Jabir Furqan, Muhammad Iqbal Abbasi, Tanveer Abbas, Rana Zulfiqar, Muhammad Wajid, Irfan Ali, Muhammad Ikhlq, Danish Sheikh and Muhammad Imran.

CHD Exome+ Consortium: This work was funded by the UK Medical Research Council (G0800270), British Heart Foundation (SP/09/002), UK National Institute for Health Research Cambridge Biomedical Research Centre, European Research Council (268834), European Commission Framework Programme 7 (HEALTH-F2-2012-279233) and Merck and Pfizer.

2. ExomeBP Consortium

Airwave: We thank all participants of the Airwave Health Monitoring Study. The study is funded by the UK Home Office, (Grant number 780-TETRA) with additional support from the National Institute for Health Research (NIHR) Imperial College Health Care NHS Trust (ICHNT) and Imperial College Biomedical Research Centre (BRC) (Grant number BRC-P38084). We thank Andy Heard and the Airwave Health Monitoring Study team for invaluable support. SNP Genotyping was performed at the Wellcome Trust Centre for Human Genetics, University of Oxford.

ASCOT: The ASCOT study was supported by Pfizer, New York, NY, USA for the ASCOT study and the collection of the ASCOT DNA repository; by Servier Research Group, Paris, France; and by Leo Laboratories, Copenhagen, Denmark. We thank all ASCOT trial participants, physicians, nurses, and practices in the participating countries for their important contribution to the study. In particular we thank Clare Muckian and David Toomey for their help in DNA extraction, storage, and handling. Genotyping of the Exome chip in ASCOT-SC and ASCOT-UK was funded by the National Institutes of Health Research (NIHR). We would also like to acknowledge the Barts and The London Genome Centre staff for genotyping the Exome chip array. This work forms part of the research programme of the NIHR Cardiovascular Biomedical Research Unit at Barts.

708 **1958BC:** We are grateful for using the British 1958 Birth Cohort DNA collection. Sample collection funded
709 by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping
710 was funded by the Wellcome Trust.

711 **BRIGHT:** This work was supported by the Medical Research Council of Great Britain (grant number
712 G9521010D), and the British Heart Foundation (grant number PG/02/128). A.F.D. was supported by the
713 British Heart Foundation (grant numbers RG/07/005/23633, SP/08/005/25115); the European Union
714 Ingenious HyperCare Consortium: Integrated Genomics, Clinical Research, and Care in Hypertension (grant
715 number LSHM-C7-2006-037093). The BRIGHT study is extremely grateful to all the patients who
716 participated in the study and the BRIGHT nursing team. The Exome chip was funded by the Wellcome Trust
717 Strategic Awards (083948 and 085475). We would also like to thank the Barts Genome Centre staff for their
718 assistance with this project. This work forms part of the research programme of the NIHR Cardiovascular
719 Biomedical Research Unit at Barts.

720 **CROATIA-Korcula:** We would like to acknowledge the invaluable contributions of the recruitment team in
721 Korcula, the administrative teams in Croatia and Edinburgh and the people of Korcula. The CROATIA-
722 Korcula study was funded by grants from the Medical Research Council (UK), European Commission
723 Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947), Ministry of Science, Education
724 and Sports of the Republic of Croatia (grant 216-1080315-0302) and the Croatian Science Foundation (grant
725 8875).

726 **DIABNORD:** We are indebted to the study participants who dedicated their time and samples to these studies.
727 We thank John Hutiaainen and Åsa Ågren (Umeå Medical Biobank) for data organization and Kerstin Enquist
728 and Thore Johansson (Västerbottens County Council) for technical assistance with DNA extraction. We also
729 thank M Sterner, M Juhas and P Storm for their expert technical assistance with genotyping and genotype data
730 preparation.

731 **EGCUT:** EGCUT received financing from European Regional Development Fund, road-map grant
732 no.3.2.0304.11-0312 and grant "Center of Excellence in Genomics" (EXCEGEN). EGCUT studies were

covered also by targeted financing from Estonian Government (IUT24-6, IUT20-60) and CTG grant (SP1GVARENG) from Development Fund of the University of Tartu.

Fenland study: The Fenland Study is funded by the Medical Research Council (MC_U106179471) and Wellcome Trust. We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams.

FINRISK 97/02: Veikko Salomaa-Dr. Salomaa was supported by the Academy of Finland, grant number 139635, and the Finnish Foundation for Cardiovascular Research.

GS:SFHS: We would like to acknowledge the invaluable contributions of the families who took part in the Generation Scotland: Scottish Family Health Study, the general practitioners and Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes academic researchers, IT staff, laboratory technicians, statisticians and research managers. SNP genotyping was performed at the Wellcome Trust Clinical Research Facility in Edinburgh. GS:SFHS is funded by the Scottish Executive Health Department, Chief Scientist Office, grant number CZD/16/6. SNP genotyping was funded by the Medical Research Council UK

GLACIER: We are indebted to the study participants who dedicated their time and samples to these studies. We thank John Hutiaainen and Åsa Ågren (Umeå Medical Biobank) for data organization and Kerstin Enquist and Thore Johansson (Västerbottens County Council) for technical assistance with DNA extraction. We also thank M Sterner, M Juhas and P Storm for their expert technical assistance with genotyping and genotype data preparation.

GoDARTS: We acknowledge the support of the Health Informatics Centre, University of Dundee for managing and supplying the anonymised data and NHS Tayside, the original data owner. We are grateful to all the participants who took part in the Go-DARTS study, to the general practitioners, to the Scottish School of Primary Care for their help in recruiting the participants, and to the whole team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses.

759 **GRAPHIC:** Recruitment and genotyping of the Genetic Regulation of Arterial Pressure of Humans in the
760 Community cohort was funded by the British Heart Foundation. N.J.S. holds a British Heart Foundation Chair
761 of Cardiology and is a senior National Institute for Health Research Investigator . This study is part of the
762 research portfolio supported by the Leicester National Institute for Health Research Biomedical Research Unit
763 in Cardiovascular Disease.

764 **HELIC-MANOLIS:** This work was funded by the Wellcome Trust (098051) and the European Research
765 Council (ERC-2011-StG 280559-SEPI). The MANOLIS study is dedicated to the memory of Manolis
766 Giannakakis, 1978–2010. We thank the residents of the Mylopotamos villages for taking part. We thank the
767 Sample Management and Genotyping Facilities staff at the Wellcome Trust Sanger Institute for sample
768 preparation, quality control and genotyping.

769 **The Nord-Trøndelag Health Study (The HUNT Study):** This is a collaboration between HUNT Research
770 Centre (Faculty of Medicine, Norwegian University of Science and Technology NTNU), Nord-Trøndelag
771 County Council, Central Norway Health Authority, and the Norwegian Institute of Public Health. CJW is
772 supported by HL094535 and HL109946.

773 **INCIPE:** The INCIPE study was supported by Foundation CARIVR, Verona, Italy and by the University of
774 Verona.

775 **LBC21 and LBC36:** We thank the cohort participants and team members who contributed to these studies.
776 Phenotype collection in the Lothian Birth Cohort 1921 was supported by the UK's Biotechnology and
777 Biological Sciences Research Council (BBSRC), The Royal Society and The Chief Scientist Office of the
778 Scottish Government. Phenotype collection in the Lothian Birth Cohort 1936 was supported by Age UK (The
779 Disconnected Mind project). Genotyping was supported by Centre for Cognitive Ageing and Cognitive
780 Epidemiology (Pilot Fund award), Age UK, and the Royal Society of Edinburgh. The work was undertaken
781 by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross
782 council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical
783 Research Council (MRC) is gratefully acknowledged.

784 **LIFELINES:** The Lifelines Cohort Study, and generation and management of GWAS genotype data for the
785 Lifelines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant
786 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of
787 Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and
788 Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University
789 Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes
790 Research Foundation. N. Verweij is supported by the Netherlands Heart Foundation (grant NHS2010B280).
791 Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-
792 generation design the health and health-related behaviours of 167,729 persons living in the North of The
793 Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-
794 demographic, behavioural, physical and psychological factors which contribute to the health and disease of
795 the general population, with a special focus on multi-morbidity and complex genetics"

796 **LOLIPOP:** The LOLIPOP study is supported by the National Institute for Health Research (NIHR)
797 Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, the British Heart
798 Foundation (SP/04/002), the Medical Research Council (G0601966, G0700931), the Wellcome Trust
799 (084723/Z/08/Z), the NIHR (RP-PG-0407-10371), European Union FP7 (EpiMigrant, 279143) and Action on
800 Hearing Loss (G51). The work was carried out in part at the NIHR/Wellcome Trust Imperial Clinical Research
801 Facility. We thank the participants and research staff who made the study possible.

802 **MDC:** This study was funded by the European Research Council (StG-282255), the Swedish Heart and Lung
803 Foundation and the Swedish Research Council. Exome chip genotyping in the MDC Cohort was supported in
804 part by NIH R01HL107816 to Kathiresan. Dr. Kathiresan's work was also funded by Fondation Leducq.

805 **NFBC1966 and NFBC1986:** The NFBC1966 and NFBC1986 studies received financial support from the
806 Academy of Finland (project grants 104781, 120315, 129269, 1114194, 24300796, Center of Excellence in
807 Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland
808 (75617), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01),
809 NIH/NIMH (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, EU
810 FP7 EurHEALTHAgeing -277849, the Medical Research Council, UK (G0500539, G0600705, G1002319,

811 G0802782, PrevMetSyn/SALVE) and the MRC, Centenary Early Career Award. The DNA extractions,
812 sample quality controls, biobank up-keeping and aliquotting was performed in the National Public Health
813 Institute, Biomedicum Helsinki, Finland and supported financially by the Academy of Finland and Biocentrum
814 Helsinki. The Section of Investigative Medicine, Imperial College is funded by grants from the MRC,
815 BBSRC, NIHR, and an Integrative Mammalian Biology (IMB) Capacity Building Award, an FP7- HEALTH-
816 2009- 241592 EuroCHIP grant and is supported by the NIHR Imperial Biomedical Research Centre Funding
817 Scheme. AIFB is also supported by the MRC, Diabetes UK, EPSRC and an EU FP7 NutriTech grant. We
818 thank the late Professor Paula Rantakallio (launch of NFBCs), and Ms Outi Tornwall and Ms Minttu Jussila
819 (DNA biobanking). The authors would like to acknowledge the contribution of the late Academician of Science
820 Leena Peltonen.

821 **OBB:** National Institute for Health Research, Oxford Biomedical Research Centre, Oxford University
822 Hospital Trust, Oxford, UK

823 **PIVUS/ULSAM:** - ULSAM and PIVUS are supported by the Swedish Research Council, Swedish Heart-
824 Lung Foundation, Swedish Diabetes Foundation and Uppsala University. The investigators want to express
825 their deepest gratitude towards the study participants.

826 **Twins UK:** -The study was funded by the Wellcome Trust: European Community's Seventh Framework
827 Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research
828 (NIHR)- funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and
829 St Thomas' NHS Foundation Trust in partnership with King's College London. SNP Genotyping was
830 performed at performed by the Genomics Core at the Wellcome Trust Centre for Human Genetics, University
831 of Oxford.

832 **UHP:** The Utrecht Health Project received grants from the Ministry of Health, Welfare and Sports (VWS),
833 the University of Utrecht, the Province of Utrecht, the Dutch Organisation of GPs and inhabitants of the
834 "Leidsche Rijn" district, Care Research, the University Medical Centre of Utrecht, and the Dutch College of
835 Healthcare Insurance Companies. We thank Utrecht, for their cooperation in this project.

836 **UKHLS:** These data are from Understanding Society: The UK Household Longitudinal Study, which is led
837 by the Institute for Social and Economic Research at the University of Essex and funded by the Economic and
838 Social Research Council. The data were collected by NatCen and the genome wide scan data were analysed
839 by the Wellcome Trust Sanger Institute. Information on how to access the data can be found on the
840 Understanding Society website <https://www.understandingsociety.ac.uk/>. The 'Understanding Society
841 Scientific Group' include the following: Understanding Society Scientific Group: Michaela Benzeval,
842 Jonathan Burton, Nicholas Buck, Annette Jäckle, Meena Kumari, Heather Laurie, Peter Lynn, Stephen
843 Pudney, Birgitta Rabe, Shamit Saggar, Noah Uhrig, Dieter Wolke

844 F. D. wishes to acknowledge the MRC Unit at the University of Bristol (MC_UU_12013/1-9)

845 A.P.M. is a Wellcome Trust Senior Research Fellow in Basic Biomedical Science (grant number WT098017).

846 M.I.M. is a Wellcome Trust Senior Investigator (WT098381); and a National Institute of Health Research
847 Senior Investigator.

848 M. H. is supported by NIH RO1 grant LM010098

849 M.D.T. has been supported by MRC fellowship G0902313.

850 F.W.A is supported by the UCL Hospitals NIHR Biomedical Research Centre and by a Dekker scholarship
851 (Junior Staff Member 2014T001) from the Dutch Heart Foundation

852 P.B.M, M.J.C, H.W.W, A. T, K.W. wish to acknowledge the NIHR Cardiovascular Biomedical Research Unit
853 at Barts and The London, Queen Mary University of London, UK for support.

854 H. Y. is funded by the European Research Council (ERC) award (323195 and SZ-50371).

855 S.R. was supported by the Academy of Finland (251217 and 255847), Center of Excellence in Complex
856 Disease Genetics, EU FP7 projects ENGAGE (201413) and BioSHaRE (261433), the Finnish Foundation for
857 Cardiovascular Research, Biocentrum Helsinki, and the Sigrid Juselius Foundation'

858 Peter Sever is an NIHR Senior Investigator and acknowledges support from the Biomedical Research Centre
859 award to Imperial College Healthcare NHS Trust.

Nicole Soranzo's research is supported by the Wellcome Trust (Grant Codes WT098051 and WT091310), the EU FP7 (EPIGENESYS Grant Code 257082 and BLUEPRINT Grant Code HEALTH-F5-2011-282510) and the National Institute for Health Research Blood and Transplant Research Unit (NIHR BTRU) in Donor Health and Genomics at the University of Cambridge in partnership with NHS Blood and Transplant (NHSBT). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or NHSBT.

3. GoT2D and T2D-GENES consortia

ADDITION: We would like to thank Torsten Lauritzen and Anneli Sandbæk for the use of the ADDITION cohort. The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen partially funded by an unrestricted donation from the Novo Nordisk Foundation (www.metabol.ku.dk).

DPS: has been financially supported by grants from the Academy of Finland (117844 and 40758, 211497, and 118590 (MU); The EVO funding of the Kuopio University Hospital from Ministry of Health and Social Affairs (5254), Finnish Funding Agency for Technology and Innovation (40058/07), Nordic Centre of Excellence on 'Systems biology in controlled dietary interventions and cohort studies, SYSDIET (070014), The Finnish Diabetes Research Foundation, Yrjö Jahnsson Foundation (56358), Sigrid Juselius Foundation and TEKES grants 70103/06 and 40058/07.

DR's EXTRA: This Study was supported by grants to **Rainer Rauramaa** by the Ministry of Education and Culture of Finland (627; 2004-2011), Academy of Finland (102318; 123885), Kuopio University Hospital, Finnish Diabetes Association, Finnish Heart Association, Päivikki and Sakari Sohlberg Foundation and by grants from European Commission FP6 Integrated Project (EXGENESIS); LSHM-CT-2004-005272, City of Kuopio and Social Insurance Institution of Finland (4/26/ 2010).

FIN-D2D 2007: This study was supported by funds from the hospital districts of Pirkanmaa; Southern Ostrobothnia; North Ostrobothnia; Central Finland and Northern Savo; the Finnish National Public Health Institute; the Finnish Diabetes Association; the Ministry of Social Affairs and Health in Finland; Finland's

886 Slottery Machine Association; the Academy of Finland [grant number 129293] and Commission of the
887 European Communities, Directorate C-Public Health [grant agreement no. 2004310].

888 **FUSION:** The FUSION study was supported by DK093757, DK072193, DK062370, and 1Z01 HG000024.

889 **Health 2006/2008:** This work was supported by the Timber Merchant Vilhelm Bang's Foundation, the
890 Danish Heart Foundation (Grant number 07-10-R61-A1754-B838-22392F), and the Health Insurance
891 Foundation (Helsefonden; Grant number 2012B233). The Health2006 was financially supported by grants
892 from the Velux Foundation; The Danish Medical Research Council, Danish Agency for Science, Technology
893 and Innovation; The Aase and Ejner Danielsens Foundation; ALK-Abello A/S, Hørsholm, Denmark, and
894 Research Centre for Prevention and Health, the Capital Region of Denmark.

895 **Inter99:** The Inter99 was initiated by Torben Jørgensen (PI), Knut Borch-Johnsen (co-PI), Hans Ibsen and
896 Troels F. Thomsen. The steering committee comprises Torben Jørgensen and Charlotta Pisinger. The study
897 was financially supported by research grants from the Danish Research Council, the Danish Centre for Health
898 Technology Assessment, Novo Nordisk Inc., Research Foundation of Copenhagen County, Ministry of
899 Internal Affairs and Health, the Danish Heart Foundation, the Danish Pharmaceutical Association, the
900 Augustinus Foundation, the Ib Henriksen Foundation, the Becket Foundation, and the Danish Diabetes
901 Association.

902 **METSIM:** The METSIM study was supported by the Academy of Finland (contract 124243), the Finnish
903 Heart Foundation, the Finnish Diabetes Foundation, Tekes (contract 1510/31/06), and the Commission of the
904 European Community (HEALTH-F2-2007 201681), and the US National Institutes of Health grants
905 DK093757, DK072193, DK062370, and 1Z01 HG000024.

906 Genotyping of the METSIM and DPS studies, and part of the FUSION study, was conducted at the Genetic
907 Resources Core Facility (GRCF) at the Johns Hopkins Institute of Genetic Medicine. The Broad Genomics
908 Platform for genotyping of the FIN-D2D 2007, FINRISK 2007, DR'sEXTRA, and FUSION studies.

909

Funding for the GoT2D and T2D-GENES studies was provided by grants NIH U01s DK085526, DK085501, DK085524, DK085545, and DK085584 (Multiethnic Study of Type 2 Diabetes Genes) and DK088389 (Low-Pass Sequencing and High-Density SNP Genotyping for Type 2 Diabetes).

C.M.L. is funded by the Wellcome Trust (086596/Z/08/Z) and the Li Ka Shing Foundation

T.T. is funded by personal grants from the Finnish Cultural Foundation, the Emil Aaltonen Foundation and the Orion-Farmos Research Foundation.

4. GTEx Project: The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health. Additional funds were provided by the NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. Donors were enrolled at Biospecimen Source Sites funded by NCI\SAIC-Frederick, Inc. (SAIC-F) subcontracts to the National Disease Research Interchange (10XS170), Roswell Park Cancer Institute 10XS171), and Science Care, Inc. (X10S172). The Laboratory, Data Analysis, and Coordinating Center (LDACC) was funded through a contract (HHSN268201000029C) to The Broad Institute, Inc. Biorepository operations were funded through an SAIC-F subcontract to Van Andel Institute (10ST1035). Additional data repository and project management were provided by SAIC-F (HHSN261200800001E). The Brain Bank was supported by a supplements to University of Miami grants DA006227, Washington University St Louis (MH101810), and the University of Pennsylvania (MH101822). The data used for the analyses described in this manuscript were obtained from the GTEx Portal and dbGaP accession number phs000424.v3.p1.

5. MuTHER consortium (TwinsUK): The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR) BioResource Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London. Tim Spector is holder of an ERC Advanced Principal Investigator award. SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR.

6. METASTROKE: provided in ²⁵

7. EchoGen consortium: provided in ²⁷

8. CHARGE-Heart Failure consortium

ARIC Study: The ARIC Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute (NHLBI) contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. CHS: This CHS research was supported by NHLBI contracts HHSN268201200036C, HSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. FHS: The FHS was supported by NHLBI (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). This work was also supported in part by grants from the NHLBI 2K24HL04334, R01HL077477, and R01HL093328 (all to RSV). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. The analyses reflect intellectual input and resource development from the FHS

investigators participating in the SNP Health Association Resource (SHARe) project. The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff of the Rotterdam Study and the participating general practitioners and pharmacists.

CL is funded by the NHLBI/NIH Contract #N01-HC-25195, NIH NIDDK R01 DK078616 and K24 DK080140, and by the Boston University School of Medicine

ATK is supported in part by NHLBI grant R01HL117078 (MAP) and NIDDK grant R01DK89256 (IBB).

Author contributions

Supervision and management of the project: JMMH and PBM. The following authors contributed to the drafting of the manuscript: JMMH, PBM, PSu, HW, ASB, FD, JPC, DRB, KW, MT, FWA, LVW, NJS, JD AKM, HY, CMM, NG, XS, TaT, DFF, MHs, OG, TF, VT. All authors critically reviewed and approved the final version of the manuscript. **Statistical analysis review:** JMMH, PSu, FD, HW, JPC, RY, NM, PBM, LVW, HY, TF, EMi, ADM, AM, AM, EE, ASB, FWA, MJC, CF, TF, SEH, ASH, JEH, JL, GM, JM, NM, APM, APo, NJS, RAS, LS, KE, MT, VT, TVV, NV, KW, AMY, WZg, NG, CML, AKM, XS, TT. **Central Data QC:** JMMH, ASB, PSu, RY, FD, HW, JPC, TF, LVW, PBM, EMi, NM, CML, NG, XS, AKM. **Central Data analysis:** JMMH, PSu, FD, HW, JPC, NG, CML, AKM, XS. **Pathway analysis and**

992 **literature review:** JMMH, DRB, PBM, MT, KW, VT, OG, AT, FWA. **GWAS lookups, eQTL analysis,**
993 **GRS, variant annotation and enrichment analyses:** JMMH, ASB, DRB, JRS, DFF, FD, MHr, PBM,
994 FWA, TT, CML, AKM, SBu. **Study Investigators in alphabetical order by consortium (CHD Exome+,**
995 **ExomeBP and GoT2D):** DSA, PA, EA, DA, ASB, RC, JD, JF, IF, PF, JWJ, FKe, ASM, SFN, BGN, DS,
996 NSa, JV, FWA, PIWB, MJB, MJC, JCC, JMC, IJD, GD, AFD, PE, TE, PWF, GG, PH, CH, KH, EI, MJ,
997 FKa, SK, JSK, LLi, MIM, OM, AMe, ADM, APM, PBM, MEN, SP, CP, OPo, DP, SR, OR, IR, VS, NJS,
998 PSe, TDS, JMS, NJW, CJW, EZ, MB, IB, FSC, LG, TH, EKH, PJ, JKu, ML, TAL, AL, KLM, HO, OPe,
999 RR, JT, MU. **Study Phenotyping in alphabetical order by consortium (CHD Exome+, ExomeBP and**
000 **GoT2D):** PA, DA, SBl, MC, JF, JWJ, FKe, KK, SFN, BGN, CJP, AR, MS, NSa, JV, WZo, RAB, MJB,
001 MJC, JCC, JMC, AFD, ASFD, LAD, TE, AF, GG, GH, PH, AS H, OLH, EI, MJ, FK, JSK, LLi, LLa, GM,
002 AMc, PM, AMe, RMg, MJN, MEN, OPo, NP, FR, VS, NJS, TDS, AVS, JMS, MT, AV, NV, NJW, TiT,
003 CC, LLH, MEJ, AK, PK, JL DPS, SM, ERBP, AS, TS, HMS, BT. **Study Data QC and analysis in**
004 **alphabetical order by consortium (CHD Exome+, ExomeBP and GoT2D):** ASB, AJMC, JMMH, JK,
005 SFN, BGN, MMN, SP, MP, PSu, ST, GV, SMW, RY, FWA, JPC, FD, AF, TF, CH, AMc, AMj, APM,
006 PBM, CP, WR, FR, NJS, MT, VT, HW, HY, NG, AKM, XS. **Exome chip data QC in alphabetical order**
007 **by consortium (CHD Exome+, ExomeBP and GoT2D):** ASB, JMMH, SFN, BGN, PSu, RY, FWA,
008 PIWB, AIFB, JCC, JPC, PD, LAD, FD, EE, CF, TF, SEH, PH, SSH, KH, JEH, EK, AMj, GM, JM, NM,
009 EMi, AMo, APM, PBM, CPN, MJN, CP, AP, WR, NRR, RAS, NS, LS, KES, MDT, VT, TVV, TVV, NV,
010 HW, HY, AMY, EZ, WZg, NG, CML, AKM, XS. **Exome chip Data analysis in alphabetical order by**
011 **consortium (CHD Exome+, ExomeBP and GoT2D):** JMMH, PSu, RY, FWA, PIWB, AIFB, RAB, MJC,
012 JCC, JPC, PD, LAD, PE, EE, CF, TF, PWF, SF, CG, SEH, PH, ASH, CH, OLH, JEH, EI, MJ, FKa, JSK,
013 DCML, LLi, JL, GM, RMr, JM, NM, MIM, PM, OM, CM, EMi, AMo, APM, RMg, PBM, CPN, MJN, TO,
014 APo, APa, WR, NRR, NJS, RAS, NS, LS, TDS, KES, MDT, ET, VT, TVV, NV, LVW, NJW, HW, HY,
015 AMY, EZ, HZ, WZg, LLB, APG, NG, MHs, JRH, AUJ, JBJ, CML, AKM, NN, XS, AS, AJS. **GRS**
016 **lookups:** AEJ, EMa, HFM, HL, HMH, JFF, MTr, RSV, WL.

017

018 **Conflict of interests**

019 N. P. has received financial support from several pharmaceutical companies that manufacture either blood
020 pressure lowering or lipid lowering agents, or both, and consultancy fees.

021 S. K. has received Research Grant-Merck, Bayer, Aegerion; SAB-Catabasis, Regeneron Genetics Center,
022 Merck, Celera; Equity-San Therapeutics, Catabasis; Consulting-Novartis, Aegerion, Bristol Myers-Squibb,
023 Sanofi, AstraZeneca, Alnylam.

024 P. Sever has received research awards from Pfizer Inc.

025 A. Malarstig and M. Uria-Nickelsen are full time employees of Pfizer.

026 D. Reily and M. Hoek are full time employees of Merck and co Inc.

027 M.J. Caulfield is Chief Scientist for Genomics England a UK Government company.

028 The authors declare no competing financial interest.

029

030 **References**

- 031 1. Lim, S.S. *et al.* A comparative risk assessment of burden of disease and injury attributable to 67 risk
032 factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden
033 of Disease Study 2010. *Lancet* **380**, 2224-60 (2012).
- 034 2. Rapsomaniki, E. *et al.* Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks,
035 healthy life-years lost, and age-specific associations in 1.25 million people. *Lancet* **383**, 1899-911
036 (2014).
- 037 3. Munroe, P.B., Barnes, M.R. & Caulfield, M.J. Advances in blood pressure genomics. *Circ Res* **112**,
038 1365-79 (2013).
- 039 4. Ehret, G.B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular
040 disease risk. *Nature* **478**, 103-109 (2011).
- 041 5. Wain, L.V. *et al.* Genome-wide association study identifies six new loci influencing pulse pressure
042 and mean arterial pressure. *Nature Genetics* (2011).
- 043 6. Johnson, T. *et al.* Blood pressure loci identified with a gene-centric array. *Am J Hum Genet* **89**, 688-
044 700 (2011).
- 045 7. Tomaszewski, M. *et al.* Genetic architecture of ambulatory blood pressure in the general population:
046 insights from cardiovascular gene-centric array. *Hypertension* **56**, 1069-76 (2010).
- 047 8. Tragante, V. *et al.* Gene-centric meta-analysis in 87,736 individuals of European ancestry identifies
048 multiple blood-pressure-related loci. *Am J Hum Genet* **94**, 349-60 (2014).
- 049 9. Ganesh, S.K. *et al.* Loci influencing blood pressure identified using a cardiovascular gene-centric
050 array. *Hum Mol Genet* (2013).
- 051 10. Simino, J. *et al.* Gene-age interactions in blood pressure regulation: a large-scale investigation with
052 the CHARGE, Global BPgen, and ICBP Consortia. *Am J Hum Genet* **95**, 24-38 (2014).
- 053 11. Zhu, X. *et al.* Meta-analysis of Correlated Traits via Summary Statistics from GWASs with an
054 Application in Hypertension. *Am J Hum Genet* **96**, 21-36 (2015).

- 055 12. Salfati, E., Morrison, A.C., Boerwinkle, E. & Chakravarti, A. Direct Estimates of the Genomic
056 Contributions to Blood Pressure Heritability within a Population-Based Cohort (ARIC). *PLoS One*
057 **10**, e0133031 (2015).
- 058 13. Schork, N.J., Murray, S.S., Frazer, K.A. & Topol, E.J. Common vs. rare allele hypotheses for
059 complex diseases. *Curr Opin Genet Dev* **19**, 212-9 (2009).
- 060 14. Nejentsev, S., Walker, N., Riches, D., Egholm, M. & Todd, J.A. Rare variants of IFIH1, a gene
061 implicated in antiviral responses, protect against type 1 diabetes. *Science* **324**, 387-9 (2009).
- 062 15. Liu, D.J. *et al.* Meta-analysis of gene-level tests for rare variant association. *Nat Genet* **46**, 200-4
063 (2014).
- 064 16. Wu, M.C. *et al.* Rare-variant association testing for sequencing data with the sequence kernel
065 association test. *Am J Hum Genet* **89**, 82-93 (2011).
- 066 17. Li, B. & Leal, S.M. Methods for detecting associations with rare variants for common diseases:
067 application to analysis of sequence data. *Am J Hum Genet* **83**, 311-21 (2008).
- 068 18. Vissers, L.E. *et al.* Heterozygous germline mutations in A2ML1 are associated with a disorder
069 clinically related to Noonan syndrome. *Eur J Hum Genet* **23**, 317-24 (2015).
- 070 19. International Consortium for Blood Pressure Genome-Wide Association, S. *et al.* Genetic variants in
071 novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103-9 (2011).
- 072 20. Coronary Artery Disease Genetics Consortium C4D. A genome-wide association study in Europeans
073 and South Asians identifies five new loci for coronary artery disease. *Nat Genet* **43**, 339-44 (2011).
- 074 21. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary
075 artery disease. *Nat Genet* **43**, 333-8 (2011).
- 076 22. Arking, D.E. *et al.* Genetic association study of QT interval highlights role for calcium signaling
077 pathways in myocardial repolarization. *Nat Genet* **46**, 826-36 (2014).
- 078 23. den Hoed, M. *et al.* Identification of heart rate-associated loci and their effects on cardiac conduction
079 and rhythm disorders. *Nat Genet* **45**, 621-31 (2013).
- 080 24. Global Lipids Genetics Consortium *et al.* Discovery and refinement of loci associated with lipid
081 levels. *Nat Genet* **45**, 1274-83 (2013).
- 082 25. Traylor, M. *et al.* Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE
083 collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol* **11**, 951-62
084 (2012).
- 085 26. Smith, N.L. *et al.* Association of genome-wide variation with the risk of incident heart failure in
086 adults of European and African ancestry: a prospective meta-analysis from the cohorts for heart and
087 aging research in genomic epidemiology (CHARGE) consortium. *Circ Cardiovasc Genet* **3**, 256-66
088 (2010).
- 089 27. Vasan, R.S. *et al.* Genetic variants associated with cardiac structure and function: a meta-analysis
090 and replication of genome-wide association data. *JAMA* **302**, 168-78 (2009).
- 091 28. Ehret, G.B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular
092 disease risk. *Nature* **478**, 103-109 (2011).
- 093 29. Russo, S.B. *et al.* Ceramide synthase 5 mediates lipid-induced autophagy and hypertrophy in
094 cardiomyocytes. *J Clin Invest* **122**, 3919-30 (2012).
- 095 30. Oudot-Mellakh, T. *et al.* Genome wide association study for plasma levels of natural anticoagulant
096 inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol* **157**, 230-9
097 (2012).
- 098 31. Smith, N.L. *et al.* Novel associations of multiple genetic loci with plasma levels of factor VII, factor
099 VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome
100 Epidemiology) Consortium. *Circulation* **121**, 1382-92 (2010).
- 101 32. Bleil, M.E., Gregorich, S.E., McConnell, D., Rosen, M.P. & Cedars, M.I. Does accelerated
102 reproductive aging underlie premenopausal risk for cardiovascular disease? *Menopause* **20**, 1139-46
103 (2013).
- 104 33. Guan, R. *et al.* rbm47, a novel RNA binding protein, regulates zebrafish head development. *Dev Dyn*
105 **242**, 1395-404 (2013).
- 106 34. Wozniak, M.A., Kwong, L., Chodniewicz, D., Klemke, R.L. & Keely, P.J. R-Ras controls membrane
107 protrusion and cell migration through the spatial regulation of Rac and Rho. *Mol Biol Cell* **16**, 84-96
108 (2005).

- 109 35. Tuckwell, D. Identification and analysis of collagen alpha 1(XXI), a novel member of the FACIT
110 collagen family. *Matrix Biol* **21**, 63-6 (2002).
- 111 36. Huang, J. *et al.* Improved imputation of low-frequency and rare variants using the UK10K haplotype
112 reference panel. *Nat Commun* **6**, 8111 (2015).
- 113 37. Segre, A.V. *et al.* Common inherited variation in mitochondrial genes is not enriched for associations
114 with type 2 diabetes or related glycemic traits. *PLoS Genet* **6**(2010).
- 115 38. Huan, T. *et al.* Integrative network analysis reveals molecular mechanisms of blood pressure
116 regulation. *Mol Syst Biol* **11**, 799 (2015).
- 117 39. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature*
118 **518**, 197-206 (2015).
- 119 40. Wood, A.R. *et al.* Defining the role of common variation in the genomic and biological architecture
120 of adult human height. *Nat Genet* **46**, 1173-86 (2014).
- 121 41. Park, H.W. *et al.* Serine-threonine kinase with-no-lysine 4 (WNK4) controls blood pressure via
122 transient receptor potential canonical 3 (TRPC3) in the vasculature. *Proc Natl Acad Sci U S A* **108**,
123 10750-5 (2011).
- 124 42. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of
125 complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).
- 126 43. Te Riet, L., van Esch, J.H., Roks, A.J., van den Meiracker, A.H. & Danser, A.H. Hypertension:
127 Renin-Angiotensin-aldosterone system alterations. *Circ Res* **116**, 960-75 (2015).
- 128 44. Gao, J. *et al.* A new strategy for treating hypertension by blocking the activity of the brain renin-
129 angiotensin system with aminopeptidase A inhibitors. *Clin Sci (Lond)* **127**, 135-48 (2014).
- 130
- 131
- 132
- 133
- 134
- 135
- 136

137 **Figure Legends**

138 **Figure 1 Study design and work flow diagram of single variant discovery analyses.** EUR=European, SAS=South
139 Asian, HIS=Hispanic, AA=African American, HTN=hypertension, BP=blood pressure, SBP=systolic blood pressure,
140 DBP= diastolic blood pressure, PP=pulse pressure, N=sample size, MAF=minor allele frequency, P = P -value
141 significance threshold, SNV=single-nucleotide variant, GWS=genome-wide significance *Further details of the selection
142 criteria are provided in the methods.

143 **Figure 2 Discovery SNV-BP associations.** Results are provided for (a) transformed SBP (b) transformed DBP (c)
144 transformed PP and (d) HTN in the European and South Asian (EUR_SAS) discovery samples. The y-axis represents –
145 $\log_{10}P$ for association. Red triangles represent variants that map to one of the 81 regions selected for replication, blue
146 triangles represent SNVs that map to previously published BP regions, and grey triangles represent all remaining SNVs.
147 SNVs are ordered according to chromosome (black lines on the outside of the plot) and physical position. Genes that
148 SNVs map to are given in the outer blocks.

149
150 **Figure 3 Overlap of the 30 novel loci associations across SBP, DBP, PP and HTN.** The Venn diagram shows which
151 of the 30 newly identified BP loci are associated with multiple BP traits. Only SNV-BP trait associations that were
152 genome-wide significant ($P < 5 \times 10^{-8}$) in the combined discovery and replication meta-analyses are listed for any given
153 BP trait, within the corresponding ancestry dataset that the given locus was validated for (see Tables 1 and 2). The
154 association of *RRAS* variant with SBP was replicated in the independent samples, but did not achieve GWS in the
155 combined discovery and replication meta-analysis and is therefore only included for SBP. HTN=hypertension,
156 SBP=systolic blood pressure, DBP= diastolic blood pressure, PP=pulse pressure.

157 **Figure 4 Study design for conditional analyses and rare variant gene-based discovery analyses.**
158 RMW=RareMetalWorker, EUR=European, SAS = South Asian, HTN=hypertension, BP=blood pressure, SBP=systolic
159 blood pressure, DBP= diastolic blood pressure, PP=pulse pressure. N=sample size, MAF=minor allele frequency, P = P -
160 value significance threshold, P_{cond} =conditional P -value significance threshold

161 **Figure 5 Locus plot for *A2ML1* and secondary amino acid structure of the gene product.** (a) Locus plot for *A2ML1*
162 associated with HTN identified through gene based tests. The variants' positions along the gene (x axis, based on

human genome build 37) and the $-\log_{10}(P\text{-value of association})$ (y axis) are indicated. The variants are colour coded: nonsense (black), missense, predicted damaging (blue), and missense (orange). The schematic above the x-axis represents the intron / exon (black vertical bars) structure, the untranslated regions are shown as grey vertical bars.

(b) The white box denotes the full-length amino acid sequence for each of the two gene products. Black numbers denote amino acid residue positions of note. Coloured boxes depict putative functional domains (see below). Coloured vertical lines indicate the amino acid substitutions corresponding to the variants depicted in the locus plots above using the same colour coding. Bold, italic indicates the SNV association with smallest *P*-value.

Dark grey – signal peptide sequence. Brown – regions of intramolecular disulfide bonds. For simplicity only those regions coinciding with variants described were indicated. Black – bait region described to interact with proteases. Purple – thiol ester sequence region aiding in interaction with proteases. Light grey – alpha helical regions thought to mediate A2ML1 interaction with LRP1, facilitating receptor-mediated endocytosis.

Table 1 Novel blood pressure trait associated loci. Variants with formal replication

Variant information				Discovery		Replication			Combined		
Locus	rsID	Chr:Pos (EA:EAF)	Trait	P_t	P_u	N	β	P	N	β	P
EUR											
<i>RNF207</i>	rs709209	1:6.28 (A:0.655)	PP	4.57×10^{-6}	1.60×10^{-6}	122,780	0.17	5.83×10^{-4}	284,683	0.20	9.62×10^{-9}
<i>C5orf56</i>	rs12521868	5:131.78 (T:0.373)	DBP	1.59×10^{-6}	3.03×10^{-7}	122,795	-0.18	2.29×10^{-5}	282,023	-0.19	6.12×10^{-11}
<i>PHACTR1</i>	rs9349379	6:12.90 (A:0.566)	SBP	2.11×10^{-8}	1.78×10^{-7}	122,809	0.24	4.06×10^{-4}	284,673	0.29	8.84×10^{-10}
<i>COL21A1</i>	rs200999181†	6:55.94 (A:0.002)	PP	3.08×10^{-8}	2.46×10^{-7}	121,487	2.70	1.90×10^{-4}	242,486	3.25	6.27×10^{-10}
<i>ABO</i>	rs687621	9:136.14 (A:0.615)	DBP	8.80×10^{-8}	2.55×10^{-7}	122,798	0.16	1.96×10^{-4}	276,014	0.19	5.45×10^{-10}
<i>ADO</i>	rs10995311	10:64.56 (C:0.567)	DBP	1.86×10^{-6}	1.14×10^{-6}	122,798	0.23	8.47×10^{-8}	266,456	0.21	1.12×10^{-12}
<i>LMO1</i>	rs110419	11:8.25 (A:0.48)	DBP	9.41×10^{-6}	2.22×10^{-5}	122,798	0.16	1.81×10^{-4}	279,935	0.16	3.04×10^{-8}
<i>OR5B12</i>	rs11229457	11:58.21 (T:0.236)	SBP	1.58×10^{-6}	4.62×10^{-5}	122,809	-0.32	7.53×10^{-5}	284,680	-0.31	2.70×10^{-8}
<i>CERS5</i>	rs7302981	12:50.54 (A:0.361)	DBP	1.35×10^{-13}	4.60×10^{-11}	122,798	0.24	2.64×10^{-8}	284,718	0.25	1.38×10^{-17}
<i>MYH6</i>	rs452036	14:23.87 (A:0.327)	PP	4.59×10^{-11}	2.80×10^{-13}	122,780	-0.21	1.81×10^{-5}	284,672	-0.28	2.96×10^{-16}
<i>DPEP1</i>	rs1126464	16:89.70 (C:0.256)	DBP	1.19×10^{-9}	4.35×10^{-11}	118,677	0.24	1.68×10^{-6}	261,564	0.28	1.02×10^{-15}
<i>TBX2</i>	rs8068318†	17:59.48 (T:0.698)	DBP	7.46×10^{-13}	5.71×10^{-10}	122,798	0.26	3.23×10^{-8}	281,978	0.26	1.95×10^{-16}
<i>RGL3</i>	rs167479	19:11.53 (T:0.486)	DBP	2.22×10^{-23}	1.97×10^{-22}	122,797	-0.29	3.01×10^{-11}	283,332	-0.33	1.99×10^{-31}
<i>PREX1</i>	rs6095241	20:47.31 (A:0.452)	DBP	5.65×10^{-6}	2.29×10^{-5}	122,798	-0.18	2.56×10^{-5}	281,322	-0.17	4.75×10^{-9}
ALL ancestry											
<i>RBM47</i>	rs35529250†	4:40.43 (T:0.01)	SBP	6.56×10^{-7}	6.15×10^{-6}	148,878	-1.43	5.02×10^{-4}	306,352	-1.55	2.42×10^{-8}
<i>OBFC1</i>	rs4387287	10:105.68 (A:0.157)	SBP	2.23×10^{-8}	1.32×10^{-7}	147,791	0.28	3.37×10^{-4}	320,494	0.36	9.12×10^{-10}
<i>RRAS</i>	rs61760904†	19:50.14 (T:0.008)	SBP	1.96×10^{-6}	1.90×10^{-5}	148,878	1.38	5.70×10^{-4}	322,664	1.50	8.45×10^{-8}

178
179
180
181
182
183
184
185
186
187

SNV-BP associations are reported for the newly identified BP loci that replicated at $P < 6.2 \times 10^{-4}$ (Bonferroni correction for the 81 variants selected for replication for a primary blood pressure trait; Methods). Loci are categorised into EUR and ALL ancestry based on the meta-analysis used to replicate the variants for the primary BP trait shown in columns labelled 'Trait'. In the columns that contains the discovery meta-analyses results, P_t represents the P -value for association of the variant with the transformed primary BP trait in the EUR_SAS discovery meta-analyses (which was also used to select the variant for replication) and P_u represents the P -value for association with the untransformed primary BP trait in the ancestry in which the variant replicated. N, β and P , which denote the number of samples, estimated allelic effect and P -value for association with the primary BP trait, are provided for the untransformed primary BP trait in the replication data and also from the combined (discovery and replication) meta-analyses. NB: ALL ancestry corresponds to all ancestries in the combined (discovery + replication) meta-analyses

Locus – Gene or region containing the SNV, rsID – dbSNP rsID. Chr:Pos (EA:EAF) – Chromosome:NCBI Build 37 position in Mb (effect allele:effect allele frequency), Trait – primary blood pressure trait for which the variant was and also replicated, β – effect estimate, N:sample size, EUR – European.

† indicates it is a non-synonymous SNV (nsSNV) or is in linkage disequilibrium with a nsSNV ($r^2 > 0.8$) that is predicted to be damaging

Table 2 Novel blood pressure trait associated loci. Variants with GWS evidence of association in combined meta-analyses

Variant information				Discovery		Replication			Combined		
Locus	rsID	Chr:Pos (EA:EAF)	Trait	P_t	P_u	N	β	P	N	β	P
EUR											
2q36.3	rs2972146	2:227.10 (T:0.652)	DBP [§] (HTN)	1.51×10^{-9}	2.47×10^{-7}	122,798	0.13	2.20×10^{-3}	275,610	0.17	8.40×10^{-9}
ZBTB38	rs16851397	3:141.13 (A:0.953)	DBP [§] (SBP)	6.87×10^{-6}	3.20×10^{-5}	122,798	-0.38	1.20×10^{-4}	284,717	-0.38	3.01×10^{-8}
PRDM6	rs1008058	5:122.44 (A:0.135)	SBP	5.09×10^{-7}	1.01×10^{-8}	43,109	0.46	3.61×10^{-3}	176,362	0.55	2.99×10^{-10}
GPR20	rs34591516	8:142.37 (T:0.055)	SBP [§] (DBP)	1.54×10^{-6}	1.01×10^{-7}	122,807	0.51	4.20×10^{-4}	282,009	0.64	6.10×10^{-10}
HOXB7	rs7406910	17:46.69 (T:0.118)	SBP	6.07×10^{-10}	2.74×10^{-9}	122,809	-0.20	4.89×10^{-2}	284,690	-0.46	3.80×10^{-8}
AMH	rs10407022†	19:2.25 (T:0.82)	PP	1.63×10^{-7}	1.73×10^{-7}	118,656	-0.19	1.62×10^{-3}	252,525	-0.26	5.94×10^{-9}
ZNF101	rs2304130	19:19.79 (A:0.914)	DBP	1.66×10^{-8}	1.92×10^{-8}	122,798	-0.17	1.71×10^{-2}	284,705	-0.29	1.53×10^{-8}
PROCR	rs867186	20:33.76 (A:0.873)	DBP	1.44×10^{-6}	4.15×10^{-7}	122,798	0.21	2.48×10^{-3}	284,722	0.26	1.19×10^{-8}
RRP1B	rs9306160	21:45.11 (T:0.374)	DBP [§] (SBP)	1.04×10^{-8}	1.90×10^{-6}	100,489	-0.16	4.30×10^{-4}	249,817	-0.18	6.80×10^{-9}
TNRC6B	rs470113	22:40.73 (A:0.804)	PP	1.48×10^{-10}	1.31×10^{-9}	122,780	-0.14	1.37×10^{-2}	284,683	-0.25	1.67×10^{-9}
ALL ancestry											
7q32.1	rs4728142	7:128.57 (A:0.433)	SBP	8.10×10^{-6}	4.21×10^{-6}	150,542	-0.21	8.62×10^{-4}	338,338	-0.24	3.45×10^{-8}
PRKAG1	rs1126930†	12:49.40 (C:0.036)	PP	2.12×10^{-6}	4.62×10^{-7}	151,481	0.36	3.74×10^{-3}	314,894	0.50	3.34×10^{-8}
SBNO1	rs1060105	12:123.81 (T:0.209)	DBP	6.66×10^{-7}	1.09×10^{-6}	150,532	-0.15	2.67×10^{-3}	336,413	-0.18	3.07×10^{-8}

SNV-BP associations are reported for the newly identified BP loci that showed genome-wide significant association ($P < 5 \times 10^{-8}$) in the combined discovery and replication meta-analyses. In the columns that contain results from the discovery meta-analyses, P_t represents the P -value for association of the variant with the transformed *primary* BP trait in the EUR_SAS discovery meta-analyses (used to select the variant for replication) and P_u represents the P -value for association with the untransformed BP trait in the ancestry in which the variant was validated. Loci are categorised into EUR and ALL ancestry based on the ancestry in which the variant showed association with a blood pressure trait at $P < 5 \times 10^{-8}$. N, β and P , which denote the number of samples, estimated allelic effect and P -value for association with the validated BP trait, are provided for the untransformed BP trait in the replication data and also from the combined (discovery and replication) meta-analyses. NB: ALL ancestry corresponds to all ancestries in the combined (discovery + replication) meta-analyses.

Locus – Gene or region containing the SNV, rsID - dbSNP rsID. Chr:Pos (EA:EAF) – Chromosome:NCBI Build 37 position in Mb (effect allele:effect allele frequency), Trait - blood pressure trait for which association is reported, EUR - European.

§ At four loci (2q36.3, ZBTB38, GPR20 and RRP1B) the primary trait used to select the variants for replication is given in parentheses because the variant associations were validated in the combined meta-analysis for the listed secondary trait. For these variants, P_t denotes the P -value for association with the primary trait, the other P -values provided are for the secondary trait.

† indicates it is a non-synonymous SNV (nsSNV) or is linkage disequilibrium with a nsSNV ($r^2 > 0.8$) that is predicted to be damaging

206 **Table 3 Results of the genetic risk score analyses across CVD traits and risk factors.**

Outcome	Units	N	DBP (per 10mmHg increase)		SBP (per 10mmHg increase)		PP (per 10mmHg increase)	
			Effect [95% CI]	P	Effect [95% CI]	P	Effect [95% CI]	P
CHD	OR	82,056	1.62 [1.28, 2.05]	5.99 x 10 ⁻⁵	1.39 [1.22, 1.59]	6.07 x 10 ⁻⁷	1.70 [1.34, 2.16]	1.20 x 10 ⁻⁵
Ischemic stroke	OR	25,799	1.93 [1.47, 2.55]	2.81 x 10 ⁻⁶	1.57 [1.35, 1.84]	1.16 x 10 ⁻⁸	2.12 [1.58, 2.84]	5.35 x 10 ⁻⁷
Cardioembolic stroke	OR	16,113	1.43 [0.86, 2.39]	0.1683	1.33 [0.99, 1.80]	0.0584	1.73 [1.00, 3.02]	0.0518
Large vessel stroke	OR	13,903	2.26 [1.25, 4.08]	0.0068	1.85 [1.32, 2.59]	3.61 x 10 ⁻⁴	3.05 [1.64, 5.68]	4.37 x 10 ⁻⁴
Small vessel stroke	OR	15,617	1.96 [1.13, 3.41]	0.0168	1.56 [1.13, 2.16]	0.0064	1.98 [1.09, 3.61]	0.0248
Heart failure	OR	13,282	1.48 [1.02, 2.17]	0.0409	1.25 [1.00, 1.57]	0.0512	1.33 [0.88, 2.02]	0.1757
Left ventricular mass	g	11,273	9.57 [3.98,15.17]	8.02 x 10 ⁻⁴	5.13 [1.77, 8.48]	0.0027	5.97 [-0.38,12.31]	0.0653
Left ventricular wall thickness	cm	11,311	0.10 [0.06, 0.13]	1.88 x 10 ⁻⁸	0.05 [0.03, 0.07]	5.52 x 10 ⁻⁶	0.05 [0.01, 0.09]	0.0187
HDL	mg/dl	80,395	0.25 [-1.00, 1.51]	0.6930	0.21 [-0.50, 0.92]	0.5622	0.47 [-0.79, 1.73]	0.4668
LDL	mg/dl	77,021	-1.57 [-5.20, 2.06]	0.3972	0.07 [-2.03, 2.16]	0.9498	1.87 [-1.86, 5.59]	0.3255
Total cholesterol	mg/dl	80,455	-1.34 [-5.90, 3.22]	0.5639	0.70 [-1.93, 3.32]	0.6029	3.68 [-0.97, 8.33]	0.1209
Triglycerides	mg/dl	77,779	0.02 [-0.03, 0.08]	0.3859	0.02 [-0.01, 0.05]	0.2697	0.03 [-0.03, 0.08]	0.3025
BMI	INVT	526,508	-0.10 [-0.18,-0.01]	0.0342	-0.07 [-0.13,-0.02]	0.0058	-0.12 [-0.23,-0.02]	0.0165
WHRadjBMI	INVT	344,369	0.03 [-0.04, 0.11]	0.4025	0.03 [-0.02, 0.08]	0.2170	0.06 [-0.03, 0.15]	0.1885
Height	INVT	458,927	0.02 [-0.15, 0.18]	0.8592	-0.04 [-0.15, 0.06]	0.4170	-0.18 [-0.37, 0.01]	0.0683
eGFR	INVT	51,039	-0.02 [-0.15, 0.11]	0.7810	-0.03 [-0.10, 0.04]	0.4080	-0.07 [-0.20, 0.06]	0.2741

207 CHD, coronary heart disease; HDL, high density lipoprotein; LDL, low density lipoprotein; eGFR, estimated glomerular filtration rate; DBP, diastolic blood pressure; SBP systolic blood pressure; PP, pulse pressure; OR, odds ratio; g, grams; INVT, inverse normally
208 transformed (hence no units); N, sample size; P, P-value of association of BP with the trait listed; CI, confidence interval. Results are considered significant if $P < 0.0038$, which corresponds to a Bonferroni correction for 13 phenotypes tested.

211 **Online Methods**

212 **Overview of discovery studies**

213 The cohorts contributing to the discovery meta-analyses comprise studies from three consortia (CHD
214 Exome+, ExomeBP, and GoT2D/T2D-GENES) with a total number of 192,763 unique samples. All
215 participants provided written informed consent and the studies were approved by their local Research Ethics
216 Committees and/or Institutional Review Boards.

217 The CHD Exome+ consortium comprised 77,385 samples: eight studies (49,898 samples) of European
218 (EUR) ancestry, two studies (27,487 samples) of South Asian (SAS) ancestry (Supplementary Table 1). The
219 ExomeBP consortium included 25 studies (75,620 samples) of EUR ancestry (Supplementary Table 1). The
220 GoT2D consortium comprised 14 studies (39,758 samples) of Northern EUR ancestry from Denmark,
221 Finland, and Sweden (Supplementary Table 1). The participating studies and their characteristics including
222 BP phenotypes are detailed in Supplementary Tables 1 and 2. Note, any studies contributing to multiple
223 consortia were only included once in all meta-analyses.

224 **Phenotypes**

225 Four blood pressure (BP) traits were analysed: systolic blood pressure (SBP), diastolic blood pressure
226 (DBP), pulse pressure (PP) and hypertension (HTN). For individuals known to be taking BP lowering
227 medication, 15/10 mmHg was added to the raw SBP/DBP values, respectively, to obtain medication-
228 adjusted SBP/DBP values⁴⁵. PP was defined as SBP minus DBP, post-adjustment. For HTN, individuals
229 were classified as hypertensive cases if they satisfied at least one of: (i) $SBP \geq 140$ mmHg, (ii) $DBP \geq 90$
230 mmHg, (iii) taking anti-hypertensive or BP lowering medication. All other individuals were included as
231 controls. The four BP traits were correlated (SBP:DBP correlations were between 0.6 and 0.8, and SBP:PP
232 correlations were ~ 0.8). However, they measure partly distinct physiological features including, cardiac
233 output, vascular resistance, and arterial stiffness, all measures for determining a cardiovascular risk profile.
234 Therefore the genetic architecture of the individual phenotypes are of interest, and a multi-phenotype
235 mapping approach was not adopted.

236

238 **Genotyping**

239 All samples were genotyped using one of the Illumina HumanExome Beadchip arrays (Supplementary Table
240 3). An Exome chip quality control Standard Operating Procedure (SOP) developed by Anubha Mahajan,
241 Neil Robertson and Will Rayner at the Wellcome Trust Centre for Human Genetics, University of Oxford
242 was used by most studies for genotype calling and QC⁴⁶ (Supplementary Table 3). All genotypes were
243 aligned to the plus strand of the human genome reference sequence (Build37) prior to any analyses and any
244 unresolved mappings were removed. Genotype cluster plots were reviewed for all the novel rare variants
245 (both lead and secondary signals) and for rare variants that contributed to the gene-based testing.

246 **Meta-analyses**

247 Meta-analyses were performed using METAL⁴⁷, for both discovery and replication analyses, using inverse
248 variance weighted fixed effect meta-analysis for the continuous traits (SBP, DBP and PP) and sample size
249 weighted meta-analysis for the binary trait (HTN).

250 **Discovery SNV analyses**

251 Analyses of both untransformed and inverse normal transformed SBP, DBP and PP were conducted within
252 each contributing study. The analyses of the transformed traits were performed in order to minimise
253 sensitivity to deviations from normality in the analysis of rare variants and for discovery of new SNV-BP
254 associations. The residuals from the null model obtained after regressing the medication-adjusted trait on the
255 covariates (age, age², sex, BMI, and disease status for CHD) within a linear regression model, were ranked
256 and inverse normalised. These normalised residuals were used to test trait-SNV associations. All SNVs that
257 passed QC were analysed for association, without any further filtering by MAF, but a minor allele count of
258 10 was used for the analysis of HTN. An additive allelic effects model was assumed.

259 Two meta-analyses were performed for each trait, one with EUR and SAS ancestries combined (EUR_SAS)
260 and another for EUR ancestry alone. Contributing studies used principal components (PCs) to adjust for
261 population stratification. Consequently minimal inflation in the association test statistics, λ , was observed (

262 $\lambda = 1.07$ for SBP, 1.10 for DBP, 1.04 for PP and <1 for HTN in the transformed discovery meta-analysis in
263 EUR_SAS; $\lambda = 1.06$ for SBP, 1.09 for DBP, 1.05 for PP and <1 for HTN in the transformed discovery
264 meta-analysis in EUR; Supplementary Figure 6). The meta-analyses were performed independently in two
265 centres and results were found to be concordant between centres. Given the studies contributing to the
266 discovery analyses were ascertained on CHD or T2D, we tested potential systematic bias in calculated effect
267 estimates amongst these studies. No evidence of bias in the overall effect estimates was obtained.

268 The results for the transformed traits were taken forward and used to select candidate SNVs for replication.
269 Results (P -values) from the transformed and untransformed analyses were strongly correlated ($r^2 > 0.9$).

270 **Replication SNV analyses**

271 SNVs associated with any of the transformed traits (SBP, DBP, PP) or HTN were annotated using the
272 Illumina SNV annotation file, humanexome-12v1_a_gene_annotation.txt, independently across two centres.
273 Given the difference in power to detect common versus low frequency and rare variant associations, two
274 different significance thresholds were chosen for SNV selection. For SNVs with $MAF \geq 0.05$, $P \leq 1 \times 10^{-5}$ was
275 selected, while, $P \leq 1 \times 10^{-4}$ was used for SNVs with $MAF < 0.05$. By choosing a significance threshold of
276 $P < 1 \times 10^{-4}$ we maximized the opportunity to follow-up rare variants (making the assumption that any true
277 signals at this threshold could replicate at Bonferroni adjusted significance, $P \leq 6.17 \times 10^{-4}$, assuming $\alpha = 0.05$
278 for 81 SNVs). All previously published BP associated SNVs and any variants in LD with them ($r^2 > 0.2$),
279 were removed from the list of associated SNVs as we aimed to replicate new findings only. SNVs for which
280 only one study contributed to the association result or showed evidence of heterogeneity ($P_{het} < 0.0001$) were
281 removed from the list as they were likely to be an artefact. Where SNVs were associated with multiple traits,
282 to minimise the number of tests performed, only the trait with the smallest P -value was selected as the
283 primary trait in which replication was sought. Where multiple SNVs fitted these selection criteria for a
284 single region, only the SNV with the smallest P -value was selected. In total, 81 SNVs were selected for
285 validation in independent samples. These 81 SNVs had concordant association results for both transformed
286 and non-transformed traits. Eighty SNVs were selected from EUR_SAS results (with consistent support in
287 EUR), and one SNV from EUR results only. In the next step, we looked up the 81 SNV-BP associations
288 using data from a separate consortium, the CHARGE+ exome chip blood pressure consortium (who had

289 analysed untransformed SBP, DBP, PP and HTN), and UHP and Lolipop (ExomeBP consortium;
290 Supplementary Tables 2 and 3). The analysed residuals from CHARGE+ were approximately normally
291 distributed in their largest studies (Supplementary Figure 7).

292 Two meta-analyses of the replication datasets were performed: one of EUR samples, and a second of EUR,
293 African American, Hispanics and SAS ancestries (“ALL”). Replication was confirmed if P (1-tailed) $<$
294 $0.05/81=6.17 \times 10^{-4}$ and the effect (beta) was in the direction observed in discovery meta-analyses for the
295 selected trait. A combined meta-analysis was performed of discovery (untransformed results as only
296 untransformed data was available from CHARGE+ exome chip blood pressure consortium) and replication
297 results across the four traits to assess the overall support for each locus. For the combined meta-analyses, a
298 GWS threshold of, $P \leq 5 \times 10^{-8}$, was used to declare a SNV as novel rather than a less stringent experiment
299 wide threshold, as GWS is used to declare significance in GWAS and we wish to minimise the possibility of
300 false positive associations. (Note that GWS is equivalent to an exome-wide threshold of $P \leq 2 \times 10^{-7}$ adjusted
301 for four traits).

302

303 Note: all validated BP-associated variants were associated at $P < 10^{-5}$ in the discovery dataset (for the primary
304 trait). Hence, we could have used the same inclusion criteria for both common and rare SNVs. Therefore the
305 optimal threshold to choose for future experiments may need further consideration.

306 **Conditional analyses and gene-based tests**

307 The RAREMETALWORKER (RMW) tool¹⁵ (version 4.13.3) that does not require individual level data to
308 perform conditional analyses and gene-based tests was used for conditional analyses. All studies that
309 contributed to the SNV discovery analyses were re-contacted and asked to run RMW. Only FENLAND,
310 GoDARTS, HELIC-MANOLIS, UKHLS and EPIC-InterAct were unable to run RMW, while two new
311 studies were included, INCIPE and NFBC1966 (Supplementary Table 1 and 2). In total, 43 studies (147,402
312 samples) were included in the EUR analyses and 45 studies (173,329 samples) in the EUR_SAS analyses
313 (Supplementary Tables 2 and 3). Comparison of discovery and RMW study level results were made
314 (Supplementary Information).

315 For each novel locus, the genomic coordinates and size of the region were defined according to
316 recombination rates (Supplementary Table 9) around the lead variant. For known loci, a 1 Mb window was
317 used (Supplementary Table 14). Conditional analyses were performed across each region, in both EUR and
318 EUR_SAS samples, for the transformed phenotype corresponding to the validated BP trait for novel loci and
319 the published BP trait for known loci.

320 Gene based tests were performed in both the EUR and EUR_SAS datasets using the Sequence Kernel
321 Association Test (SKAT)¹⁶ method implemented in RMW as it allows for the SNVs to have different
322 directions and magnitudes of effect. Burden tests were also performed but are not presented as only SKAT
323 provided significant results. The variants in the gene-based tests using SKAT were weighted using the
324 default settings, *i.e.* a beta distribution density function to up-weight rare variants, $\text{Beta}(\text{MAF}_j, 1, 25)$ where
325 MAF_j represents the pooled MAF for variant j across all studies. Analyses were restricted to coding SNVs
326 with $\text{MAF} < 5\%$ and $< 1\%$. Genes were deemed to be associated if $P < 2.8 \times 10^{-6}$ (Bonferroni adjusted for
327 17,996 genes). To confirm the gene associations were not attributable to a solitary SNV, a gene-based test
328 conditional on the most associated SNV was performed ($P_{\text{conditional}} < 0.001$). The QC of all SNVs
329 contributing to the gene based tests including the number of samples and studies were checked prior to
330 claiming association. We sought replication of associated genes in the CHARGE+ exome chip blood
331 pressure consortium.

332

333 **Pathway analyses with MAGENTA**

334 We tested seven databases in MAGENTA³⁷ (BioCarta, Kyoto Encyclopedia of Genes and Genomes,
335 Ingenuity, Panther, Panther Biological Processes, Panther Molecular Functions and Reactome) for
336 overrepresentation of the SNV discovery results from both EUR and EUR_SAS ancestries. Each of the four
337 BP phenotypes were tested. Pathways exhibiting $P < 0.01$ and $\text{FDR} < 5\%$ were considered statistically
338 significant.

339 **GeneGo MetaCore Network analyses**

340 A set of BP genes based on previously published studies and our current results (locus defined as $r^2 > 0.4$ and
 341 500kb on either side of the lead SNV; Supplementary Table 19) were tested for enrichment using the
 342 THOMSON REUTERS MetaCore™ Single Experiment Analysis workflow tool. The data were mapped
 343 onto selected MetaCore ontology databases: pathway maps, process networks, GO processes and diseases /
 344 biomarkers, for which functional information is derived from experimental literature. Outputs were sorted
 345 based on P - and FDR-values. A gene set was considered enriched for a particular process if $P < 0.05$ and
 346 FDR < 5%.

347 **Genetic Risk Score** 348 349

350 To assess the effect of BP on CHD, ischemic stroke (and subtypes: large vessel, small vessel and
 351 cardioembolic stroke) left ventricular mass, left ventricular wall thickness, heart failure, HDL-c, LDL-c,
 352 total cholesterol, triglycerides and eGFR, we performed a weighted generalized linear regression of the
 353 genetic associations with each outcome variable on the genetic associations with BP.
 354 When genetic variants are uncorrelated, the estimates from such a weighted linear regression analysis using
 355 summarized data, and a genetic risk score analysis using individual-level data, are equal⁴⁸. We refer to the
 356 analysis as a genetic risk score (also known as a polygenic risk score) analysis as this is likely to be more
 357 familiar to applied readers. As some of the genetic variants in our analysis are correlated, a generalized
 358 weighted linear regression model is fitted that accounts for the correlations between variants, as follows:
 359 If β_X are the genetic associations (beta-coefficients) with the risk factor (here, BP) and β_Y are the genetic
 360 associations with the outcome, then the causal estimate from a weighted generalized linear regression is $(\beta_X^T$
 361 $\Omega^{-1} \beta_X)^{-1} \beta_X^T \Omega^{-1} \beta_Y$, with standard error,

$$362 \quad \hat{\sigma} \sqrt{(\beta_X^T \Omega^{-1} \beta_X)^{-1}},$$

363 where T is a matrix transpose, $\hat{\sigma}$ is the estimate of the residual standard error from the regression model, and
 364 the weighting matrix Ω has terms

$$365 \quad \Omega_{j_1 j_2} = \sigma_{Y j_1} \sigma_{Y j_2} \rho_{j_1 j_2}$$

366 , where $\sigma_{Y j}$ is the standard error of the genetic association with the outcome for the j th SNV, and $\rho_{j_1 j_2}$ is the
 367 correlation between the j_1 th and j_2 th SNVs. The presence of the estimated residual standard error allows for

368 heterogeneity between the causal estimates from the individual SNVs as overdispersion in the regression
369 model (in the case of underdispersion, the residual standard error estimate is set to unity). This is equivalent
370 to combining the causal estimates from each SNV using a multiplicative random-effects model⁴⁹.

371

372 For each of SBP, DBP and PP, the score was created using both the novel and known BP SNVs or a close
373 proxy ($r^2>0.8$). Both the sentinel SNV association and any secondary SNV associations that remained after
374 adjusting for the sentinel SNV were included in the genetic risk score. For the 30 validated novel SNV-BP
375 associations, β s were taken from the independent replication analyses (Table 1 and 2) to weight the SNV in
376 the genetic risk score. For the secondary SNVs from the seven novel loci and five known loci, β s were taken
377 from the discovery analyses (Supplementary Tables 10 and 15). For the 82 known SNVs, 43 were either
378 genotyped or had proxies on the Exome chip and the β s were taken from discovery results (Supplementary
379 Table 13), the remaining β s were taken from published effect estimates. This strategy for selecting betas for
380 use in the GRS was taken to minimize the influence of winner's curse. The associations between the BP
381 variants with CHD, HDL-c, LDL-c, total cholesterol, log(triglycerides) and log(eGFR) were obtained using
382 the CHD Exome+ Consortium studies, the associations with BMI, waist-hip ratio adjusted BMI and height
383 from the GIANT consortium (unpublished data), ischemic stroke from METASTROKE²⁵, and left
384 ventricular mass, left ventricular wall thickness and heart failure from EchoGen²⁷ and CHARGE-HF²⁶. A
385 causal interpretation of the association of GRS with the outcome as the effect of BP on the outcome assumes
386 that the effects of genetic variants on the outcome are mediated via blood pressure and not via alternate
387 causal pathways, for example via LV thickness. There are also limitations of the Mendelian randomization
388 approach in distinguishing between the causal effects of different measures of blood pressure, due to the
389 paucity of genetic variants associated with only one measure of blood pressure.

390

391 **eQTL analyses**

392 The MuTHER dataset contains gene expression data from 850 UK twins for 23,596 probes and 2,029,988
393 (HapMap 2 imputed) SNVs. All cis-associated SNVs with FDR<1%, within each of the 30 novel regions
394 (IMPUTE info score >0.8) were extracted from the MuTHER project dataset for, LCL (n=777), adipose

(n=776) and skin (n=667)⁵⁰. The pilot phase of the GTEx Project (dbGaP Accession phs000424.v3.p1) provides expression data from up to 156 individuals for 52,576 genes and 6,820,472 genotyped SNVs (imputed to 1000 Genomes project, MAF \geq 5%)⁵¹. The eQTL analysis was focused on subcutaneous adipose tissue (n=94), tibial artery (n=112), heart (left ventricle) (n=83), lung (n=119), skeletal muscle (n=138), tibial nerve (n=88), skin (sun exposed, lower leg) (n=96), thyroid (n=105) and whole blood (n=156) which have >80 samples and genes expressed at least 0.1 RPKM in 10 or more individuals in a given tissue. All transcripts with a transcription start site (TSS) within one of the 30 new BP loci and for which there was a cis-associated SNV (IMPUTE info score >0.4) within 1Mb of the TSS at FDR<5%, were identified. Kidney was not evaluated because the sample size was too small (n=8). From each resource, we report eQTL signals, which reach the resource-specific thresholds for significance described above, for SNVs that are in LD (r^2 >0.8) with our sentinel SNV.

For identified eQTLs, we tested whether they colocalised with the BP associated SNV⁵². Colocalisation analyses were considered to be significant if the posterior probability of colocalisation was greater than 0.95.

Annotation of variants

In silico prediction of the functional effect of associated variants was based on the annotation from dbSNP, the Ensembl Variant Effect Predictor tool and the Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA.

Trait variance explained

The percentage trait variance explained for SBP, DBP, PP was assessed with 5,861 individuals with complete information for all phenotypes and covariates from the population-based cohort, 1958BC.

Two genetic models were investigated: one containing the 43 previously known BP associated SNVs covered on the Exome chip; the other additionally including the 30 novel lead SNVs and 9 conditionally independent SNVs from both novel and known loci. These nine conditionally independent SNVs were taken from the EUR results, as 1958BC is EUR. They included four from novel loci (*PREX1*, *COL21A1*, *PRKAG1* and *MYH6* (there was only 1 in EUR); Supplementary Table 10) and five from known loci (*ST7L-CAPZA1-MOV10*, *FIGN-GRB14*, *ENPEP*, *TBX5-TBX3* and *HOXC4*; Supplementary Table 15).

The residual trait was obtained by adjusting each of the BP traits in a regression model with sex and BMI variables (not age or age² as all 1958BC individuals were aged 44 years). The residual trait was regressed on all SNVs within the corresponding model and adjusted for the first ten PCs. The R² calculated from this regression model was used as the percentage trait variance explained.

Monogenic Enrichment analyses

To determine if sub-significant signals of association were present in a set of genes associated with monogenic forms of disease, we performed an enrichment analysis of the discovery single variant meta-analyses association results for all four traits, both for EUR and EUR_SAS datasets.

The monogenic gene set included: *WNK1*, *WNK4*, *KLHL3*, *CUL3*, *PPARG*, *NR3C2*, *CYP11B1*, *CYP11B2*, *CYP17A1*, *HSD11B2*, *SCNN1A*, *SCNN1B*, *SCNN1G*, *CLCNKB*, *KCNJ1*, *SLC12A1*, *SLC12A3*³. The association results of coding SNVs in these genes were extracted and the number of tests with $P < 0.001$ observed. In order to determine how often such an observation would be observed by chance, we constructed 1,000 matched gene sets. The matching criteria for each monogenic gene was the intersection of all genes in the same exon length quintile and all genes in the same coding variant count decile. Within the matched sets, the number of variants with $P < 0.001$ was observed. The empirical P -value was calculated as the fraction of matched sets with an equal or larger number of variants less than 0.001.

References

45. Tobin, M.D., Sheehan, N.A., Scurrah, K.J. & Burton, P.R. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med* **24**, 2911-35 (2005).
46. Mahajan, A. *et al.* Identification and Functional Characterization of G6PC2 Coding Variants Influencing Glycemic Traits Define an Effector Transcript at the G6PC2-ABCB11 Locus. *PLoS Genet* **11**, e1004876 (2015).
47. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
48. Burgess, S. & Thompson, S.G. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* **181**, 251-60 (2015).
49. Thompson, S.G. & Sharp, S.J. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med* **18**, 2693-708 (1999).

- 452 50. Nica, A.C. *et al.* The architecture of gene regulatory variation across multiple human tissues: the
453 MuTHER study. *PLoS Genet* **7**, e1002003 (2011).
- 454 51. GTEx, Consortium,. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis:
455 multitissue gene regulation in humans. *Science* **348**, 648-60 (2015).
- 456 52. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association
457 studies using summary statistics. *PLoS Genet* **10**, e1004383 (2014).
- 458